PITMAN CLOSENESS OF MAXIMUM LIKELIHOOD ESTIMATORS UNDER TYPE-II HYBRID CENSORING WITH EXPONENTIAL LIFETIMES

PITMAN CLOSENESS OF MAXIMUM LIKELIHOOD ESTIMATORS UNDER TYPE-II HYBRID CENSORING WITH EXPONENTIAL LIFETIMES

By ANNA LY, HBSc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2025)

Hamilton, Ontario, Canada (Mathematics and Statistics)

TITLE:	Pitman Closeness of Maximum Likelihood Estimators under Type-II Hybrid Censoring with Exponential Lifetimes
AUTHOR:	Anna Ly HBSc (Applied Statistics), University of Toronto, Mississauga, Canada
SUPERVISOR:	Dr. Katherine Davies
NUMBER OF PAGES:	xiii, 125

Lay Abstract

Researchers are often interested in the time it takes for a certain event to happen. For example, in medical studies, we may ask how long it takes a patient to recover, while in engineering, we may study how long a product works before it fails. This type of information, which measures the time until an event occurs, is called lifetime data. Collecting such data can be difficult because studies often end before every recovery or failure has been observed, resulting in incomplete data.

To make sense of incomplete data, statisticians use statistical inference, a process where they make inferences about the population from available data. There is a special type of statistical inference, called estimation, where mathematical formulas called estimators are used to approximate important features of said population.

This thesis examines how to decide which estimator is more accurate among a given class under a specific data collection scheme. Using a mathematical tool called the Pitman closeness criterion, we derive and compute exact expressions for making pairwise comparisons among three different estimators that depend on the length of the study and the number of observations collected. Our results, based on this criterion, support the intuitive idea that extending the study period or increasing the number of observations leads to producing a better estimator according to the Pitman closeness criterion in a particular data collection scheme.

Abstract

The Pitman closeness (PC) criterion is a method to compare two statistical estimators. Assuming that the lifetime data follow an exponential distribution with scale parameter θ, prior work had computed the PC probabilities for estimators of θ based on Type-I right-censoring, Type-II right-censoring and Type-I hybrid censoring schemes (HCS). However, the derivation of the PC under a Type-II HCS has not yet been addressed in the literature.

This thesis examines two comparisons of maximum likelihood estimators for θ, the scale parameter, for exponentially distributed lifetimes arising from the Type-II HCS: (1) between estimators corresponding to different numbers of observed failures, and (2) between estimators with different censoring times. Closed-form expressions for the PC probabilities are derived, and numerical results are reported for various sample sizes, censoring times, and study durations. Numerical results show that increasing the pre-fixed termination time or the number of failures led to an estimator that was always Pitman closer to the true parameter. These findings confirm the intuition that increasing the termination time or the number of observed failures will usually lead to an estimator that is Pitman closer than one based on a shorter termination time or fewer observed failures.

Acknowledgements

Sitting on the couch of my first-year residence dorm, someone important to me suggested that I should do a master's degree.

"I could never do it," I responded. "I'm not smart enough."

"Well, I'm going to," he said. "You're hardworking. You should join me."

Ironically, out of the two of us, only I am pursuing graduate school. We do not talk anymore, but I appreciate that he was the first one who believed in me.

When I was in high school, one of my teachers said the only time they cried was when they were a master's student. Although graduate school has been admittedly very difficult, I do not share the same sentiment that the process will break someone into tears. Perhaps this could be attributed to those around me.

I would like to give a sincere thanks to my graduate school sweetheart, Simran, for supporting me throughout this journey. I appreciate the car rides to school every morning, his humour, and his undying support, even though I am often difficult to deal with. I love that we complement each other well; I'm happy to help him with code, and likewise, he helps me deal with some theoretical problems. We truly are a power couple.

Furthermore, I am grateful to Professor Katherine Davies for her support throughout my entire thesis. I appreciate her patience while I was occupied with coursework, her guidance when I was stuck on a derivation, and her thorough re-reads of my thesis to ensure everything is perfect. Her reference letter was especially well-written, which helped me obtain a lecturing position at UTM.

I must give thanks to the examining committee members. I am truly appreciative of Professor Balakrishnan for taking his time to assist me with the derivations. He is even so kind to help my friends and I when we bumped into him at Annalakshmi! Additionally, thanks to Professor Feng for his humourous lectures and wisdom. His stories and background are so fascinating; I could listen to them all day.

Moreover, I am very much obliged to my parents for their support. I'm glad that they accepted my aspirations in academia and have helped me plenty with my yearly moves all around Ontario for the past several years. I will likely make another annoying move for the PhD. I also want to shout out to my brother for driving me around anywhere I want every time I visit family in Niagara Falls.

In addition, I want to show my appreciation to Professor Luai Al Labadi. Out of every professor I've met, he encouraged me to pursue graduate studies the most! He gave me ample support, opportunities, and knowledge, which prepared me for graduate school. Even now, I enjoy that we still keep in touch, and I hope to teach alongside him one day.

It is regrettable that I was too occupied to engage with the graduate mathematics and statistics community deeply. My short interactions with the members of Math & Statistics Graduate Society (MSGS) have been lovely. I love that at McMaster, we graduate students treat each other with respect and kindness instead of hostility.

Table of Contents

List of Figures

List of Tables

Notation and Abbreviations

Notation

$\lvert \cdot \rvert$	Denotes the absolute value of the argument, where "$\cdot$" serves as a placeholder for a real number, defined as appropriate in context.
$\mathbb{P}(\cdot)$	The probability of the placeholder "$\cdot$" occurring, which is precisely defined as needed.
$X \sim \mathcal{D}(\cdot)$	The random variable X follows a placeholder $\mathcal{D}$ distribution. $\mathcal{D}$ and "$\cdot$" are stated when relevant.
$\mathit{Exponential}(\theta)$	The exponential distribution with rate parameter θ.
$f_X(x)$	The probability density function (pdf) of a random variable X evaluated at x.
$F_X(x)$	The cumulative distribution function (cdf) of a random variable X evaluated at x.
$f_X(x; \cdot)$	An alternative notation for the pdf of a random variable X evaluated at x, where "$\cdot$" serves as a placeholder for the parameters for the distribution.

$F_X(x;\cdot)$	An alternative notation for the cdf of a random variable X evaluated at x, where "$\cdot$" serves as a placeholder for the parameters for the distribution.
$L(\cdot)$	The likelihood function where "$\cdot$" serves as a placeholder for the parameter vector for the distribution.
$X_{i:n}$	The ith order statistic from a sample of size n.
$\mathbf{1}_{[a,b]}(x)$	An indicator function where we have 1 if $a \leq x \leq b$ and 0 otherwise.

Abbreviations

PC	Pitman closeness criterion.
HCS	Hybrid censoring scheme.
pdf	Probability density function.
cdf	Cumulative distribution function.
i.i.d.	Independently and identically distributed.
MLE	Maximum likelihood estimator.
UMVUE	Unique minimum variance unbiased estimator.
MSE	Mean square error.
BLUE	Best linear unbiased estimator.
BLIE	Best linear invariant estimator.

Chapter 1

Introduction

Lifetime experiments, which aim to measure the time until an event occurs, are commonly employed in healthcare and engineering settings. For example, researchers measure how long it takes for patients to recover from a disease, or engineers test how long their product can be used for. There are many costs and time considerations when conducting lifetime experiments, which may result in incomplete data, meaning data for which the event of interest (e.g., recovery, failure) has not been observed for all subjects. For instance, patients can pass away prematurely, or a product may last longer than the study time.

Most experiments are conducted under right-censoring, where only a lower bound on lifetimes are fully observed for the observations, often using either Type-I or Type-II censoring schemes. However, there are hybrid versions, such as the Type-I and Type-II Hybrid Censoring Schemes (HCS), which may provide more information for estimating parameters for lifetime distributions. These naming conventions (Type-I HCS, Type-II HCS) were first introduced by Childs et al. [11].

Because these schemes affect how much information is available, it is common when analysing lifetime data to model the times to event using a specific probability distribution. In this context, these are often referred to as lifetime distributions, although functionally they are standard probability distributions applied to lifetime data. Some recent examples in the literature use an exponential distribution to analyse electronic components [12] and fluorescence decay [24]. Other common distributions include the log-normal, log-logistic, gamma, inverse Gaussian, and the Weibull distribution [19].

Previous studies have applied the Pitman closeness (PC) criterion, a method to compare two estimators, to censored samples drawn from the exponential distribution. These include investigations under Type-I censoring, Type-II censoring, and Type-I HCS [7, 8, 13]. The work about Type-I HCS specifically examined how increasing the pre-specified number of observed events r or the censoring time T influences the closeness of an estimator to the value of the scale parameter from the exponential distribution under the PC criterion. However, an analogous comparison has not yet been carried out.

This thesis focuses on the derivation of the PC probabilities among maximum likelihood estimators (MLEs) of the scale parameter θ obtained under Type-II hybrid censoring schemes from an exponential lifetime distribution. The result provides a suggestion for researchers designing experiments under a Type-II hybrid censoring scheme. Specifically, this determines whether increasing the number of observed failures or extending the total data collection time leads to an estimator that is Pitman closer to the true parameter, which can be used as a heuristic to identify a more accurate estimator.

The rest of this thesis is organised as follows. Chapter 2 introduces key preliminary concepts in detail, including Pitman closeness, lifetime data, and various censoring schemes. Chapters 3 and 4 present comparisons between MLEs of the scale parameter θ under a Type-II HCS from a lifetime exponential distribution. Chapter 3 examines how increasing the number of observed failures affects estimator performance, while Chapter 4 investigates the impact of extending the total data collection time. Next, we present numerical results from computing the PC probability for varying cases in Chapter 5. Finally, we make our concluding remarks and suggest new avenues for future research in Chapter 6.

Chapter 2

Preliminaries

In this chapter, we present relevant preliminaries to understand the method used for comparing estimators, as well as the structure and properties of the exponential distribution.

2.1 Lifetime Data Analysis

Lifetime refers to the length of time from a defined starting point (such as the beginning of observation or the start of product use) until a specified event, often called a **failure**, occurs. The definition of "failure" depends on the situation and does not imply something negative. For example, in engineering, this might mean a product no longer works properly. In healthcare, it may be that a patient has been cured or has passed away.

Lifetime data, as the name suggests, consists of observations of these lifetimes. This type of data can also be referred to as **survival** or **failure time** data, depending on the context. Such data arise in biomedical sciences, epidemiology, engineering, reliability studies, and many other fields. For instance, in healthcare, one might record the time it takes for a patient to recover from an illness to estimate average recovery times. In manufacturing, lifetime data are collected in reliability studies to assess how long a product lasts before it stops functioning as intended.

When analysing lifetime data, it is typical to assume an underlying probability distribution describing the time to event. In the context of lifetime data analysis, such models are commonly termed as **lifetime distributions**; however, they are mathematically identical to conventional probability distributions.

Lifetime distributions can be either continuous or discrete. Continuing the healthcare example, one might measure recovery time in days, which is discrete, or in hours or minutes, which can be treated as continuous. This thesis, however, focuses on continuous lifetime distributions; consequently, all notations, including cumulative distribution functions, will be discussed in their continuous form.

Let T be a non-negative random variable that represents the lifetime of a subject under study, which again is the duration until an event of interest occurs. Let $f_T(t)$ denote the probability density function (pdf) of T; usually, $f_T(t) \in [0, \infty)$. Also, let $F_T(t)$ represent the cumulative density function (cdf) of T. We can write the cdf as:

$$F_T(t) = \mathbb{P}(T \leq t) = \int_0^t f_T(x)dx. \tag{2.1.1}$$

Here, the cdf represents the probability that a lifetime ends before time t. It is common for researchers to be more interested in the probability that a lifetime lasts beyond time t. Hence, the **survival function**, sometimes referred to as the

reliability function, denoted as $S_T(t)$, is shown below,

$$S_T(t) = \mathbb{P}(T \geq t) = \int_t^{\infty} f_T(x)dx, \tag{2.1.2}$$

and this function represents the probability that an individual or object survives beyond a specified time point t.

2.1.1 Lifetime Distributions

When analysing lifetime data, researchers often rely on well-known probability distributions such as the log-normal, log-logistic, gamma, inverse Gaussian, and, most commonly, the Weibull distribution [19].

The pdf for the Weibull distribution is defined as:

$$f_X(x) = \frac{k}{\theta}\left(\frac{x}{\theta}\right)^{k-1} \exp\left\{-\left(\frac{x}{\theta}\right)^{k}\right\}, \quad x \geq 0,\, k > 0,\, \theta > 0, \tag{2.1.3}$$

where k is the shape parameter, and θ is the scale parameter. A special case of the Weibull distribution is the exponential distribution, which occurs when $k = 1$. That is, assuming we have the same scale parameter θ, the pdf of an **exponential distribution** is given by:

$$f_X(x) = \frac{1}{\theta}e^{-x/\theta}, \quad x \geq 0,\, \theta > 0. \tag{2.1.4}$$

The corresponding cdf is:

$$F_X(x) = 1 - e^{-x/\theta}, \quad x \geq 0,\, \theta > 0. \tag{2.1.5}$$

Furthermore, if there exists $\epsilon > 0$ such that $\mathbb{E}[e^{Xt}]$ is finite for all $t \in (-\epsilon, \epsilon)$, the moment generating function of the exponential distribution is:

$$M_X(t) = \mathbb{E}[e^{Xt}] = (1 - \theta t)^{-1}, \;\; t \neq \theta^{-1}. \tag{2.1.6}$$

The exponential distribution has nice properties. For instance, the exponential distribution is scale-invariant under reparameterisation. That is, let $X \sim$ *Exponential*(θ), where θ represents the scale parameter, and let $Y = X/\theta$. Then,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(X/\theta \leq y) = \mathbb{P}(X \leq \theta y) = F_X(\theta y), \\ f_Y(y) &= f_X(\theta y)\theta = \frac{1}{\theta} e^{-(\theta y)/\theta}\theta = e^{-y}. \end{aligned} \tag{2.1.7}$$

Hence, $Y \sim$ *Exponential*(1). This result shows us that when working with estimators for the scale parameter of an exponential distribution, results can be rescaled to apply to any value of θ.

Furthermore, there is an interesting relationship between the exponential distribution and the gamma distribution. Let $V \sim$ *Gamma*(α, β), where α represents the shape and β represents the scale. The shape and scale parameterisation of the gamma pdf is defined as:

$$f_V(v; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} v^{\alpha-1} e^{-v/\beta}, \quad x \geq 0, \, \alpha > 0, \, \beta > 0. \tag{2.1.8}$$

Similarly, if $\mathbb{E}[e^{Vt}]$ is finite for $\epsilon > 0, \;\; t \in (-\epsilon, \epsilon)$, then we define the moment generating function of the gamma distribution to be:

$$M_V(t; \alpha, \beta) = \mathbb{E}[e^{Vt}] = (1 - \beta t)^{-\alpha}. \tag{2.1.9}$$

The gamma distribution has the useful property that the sum of independent exponential random variables with a common scale parameter θ follows a gamma distribution. A formal proof is given in Theorem (2.1.18) in Durrett's textbook [14], and we restate the lemma below for reference.

Lemma 2.1.1. *Let X_1, X_2, $\ldots$, X_n be be independent and identically distributed (i.i.d.) distributed as Exponential(θ). Then, the distribution of $\xi := X_1 + X_2 + \cdots + X_n$ is Gamma(n, θ) where the first value (n) represents the shape parameter and the second value (θ) represents the scale parameter.*

2.1.2 Order Statistics

Order statistics, as their name suggests, are defined by sorting a random sample in increasing order. That is, given we have a random sample of size n: $X_1, X_2, \ldots, X_n$, we denote the corresponding order statistics as $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ where $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$. Order statistics appear in numerous statistical contexts; here, we consider their application specifically to lifetime data.

The pdf of the ith order statistic, as derived by Barry et al. [3] as Equation (2.2.2), is given in the following lemma:

Lemma 2.1.2. *Let X_1, X_2, $\ldots$, X_n be i.i.d. random variables with probability density function $f_X(x)$ and cumulative distribution function $F_X(x)$. Denote their order statistics by $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$. The probability density function of $X_{i:n}$, for $i \in \{1, \ldots, n\}$, is:*

$$f_{X_{i:n}}(x) = \frac{n!}{(i-1)!(n-i)!}\{F_X(x)\}^{i-1}\{1 - F_X(x)\}^{n-i} f_X(x), \ \ -\infty < x < \infty. \tag{2.1.10}$$

There is an interesting result involving spacings of exponential distributions derived as Theorem (4.6.1) in Barry et al., [3]:

Theorem 2.1.1. *Let* $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ *be the order statistics from the standard exponential distribution from a sample of size* n*. We will set* $X_{0:n} = 0$*. Then consider the random variable of the form:*

$$Z_i = (n - i + 1)(X_{i:n} - X_{i-1:n}), \quad i \in \{1, 2, \ldots, n\}. \tag{2.1.11}$$

We have that $Z_1, Z_2, \ldots, Z_n$ *are all statistically independent and have standard exponential distributions.*

This result may seem overly specific, but it frequently arises in the context of order statistics from the exponential distribution.

Order statistics are central to constructing **L-estimators**, which are linear combinations of order statistics. Common examples include the median, minimum, maximum, and quantiles, all of which are generally robust to outliers. A detailed treatment of L-statistics can be found in Serfling's book [23].

Beyond their role in descriptive statistics, some types of data are inherently ordered. This is especially relevant in life-testing experiments with missing or incomplete data, also known as censored data, a topic explored in the next subchapter.

2.1.3 Standard Right-Censoring Schemes

Many unexpected events may occur during the data collection phase for lifetime data. Morbid events, such as patients passing away before they can recover, are not unique. Furthermore, sometimes experiments can take longer than the original time allocated, resulting in experiments that may be prematurely terminated. Hence, we discuss the notion of **censored data**, a broad term that implies incomplete observations. There are specific types of censored data; we will focus on discussing extensions of right-censored data.

Right censoring refers to experiments where only a lower bound on the lifetime is observed. This means that for some subjects, the event of interest may occur later than the observation period, and the exact time is unknown. For example, in a healthcare study tracking patient recovery, a patient may recover soon after the observation period ends or potentially much later; the precise recovery time remains uncertain. It is common to exclude the suffix "right" when describing right-censored experiments, and we will adopt this convention throughout. There are different types of right censoring schemes, such as Type-I, Type-II, Progressive Type-I, Progressive Type-II, Type-I Hybrid, and Type-II Hybrid.

Type-I refers to situations where the experiment ends at a predetermined time, thus the number of observed failures is random. On the other hand, **Type-II** occurs when the experiment continues until a pre-determined number of events have occurred, while the termination time is random. In summary, Type-I is fixed by time, and Type-II is fixed by the event count. A good comprehensive introduction to lifetime data and right-censoring schemes is discussed in Lawless' textbook [19].

There are modifications to Type-I and Type-II censoring schemes to add an extra layer of complexity. An example that we will briefly discuss is Progressive Type-I and Progressive Type-II, where censoring takes place progressively at r different stages.

In a lifetime experiment, any unit that has not yet failed is referred to as a **live unit** or **surviving unit**. Progressive Type-I censoring describes an experimental setup where a set of termination times, $t_1, t_2, \ldots, t_k$, is fixed in advance, and at each of these times a predetermined number of live units $m_1, m_2, \ldots, m_k$ are randomly removed from the study.

In contrast, **Progressive Type-II** censoring is both more commonly studied and more widely applied than Progressive Type-I. It involves removing a predetermined number of live units at the time of certain failures. Suppose an experiment begins with n units, and the goal is to observe exactly k failures. Let $m_1, m_2, \ldots, m_k$ denote the numbers of live units to be removed in addition to the observations that fail at the times of the 1st, 2nd, ..., kth failures. For instance, when the first failure occurs, an additional m_1 of the remaining $n-1$ live units are withdrawn at random. At the second failure, m_2 units are removed from the remaining $n-2-m_1$ live units, and this process continues until $n - r - m_1 - m_2 - \cdots - m_k$ units remain.

Additional references on Progressive Type-I and Progressive Type-II censoring are provided in books written by Balakrishnan & Aggarwala [4] and Balakrishnan & Cramer [6].

There are some issues with Type-I and Type-II censoring schemes. For example, in Type-I censoring schemes, if the termination time T is small, it is possible to observe no failures. Hence, it would be impossible to make any statistical inferences.

Other flaws regarding Type-I and Type-II censoring schemes are that they do not account for both the number of observations as well as the termination time, although realistically, most researchers have an idea of both factors in the planning phase. For instance, in Type-I censoring, if the termination time T is large, then there is a possibility that the number of observed failures r will exceed the number of failures that the experimenter had planned to account for, meaning the test could be unnecessarily prolonged.

Furthermore, long experiments are often costly and impractical. In engineering, for instance, testing the durability of mobile phones over many years is unrealistic when new models are released annually. Similarly, in healthcare, it is difficult for participants to remain in long-term studies without compensation, and personal circumstances may cause them to withdraw.

Conversely, in Type-II censoring, since the number of planned observations r is fixed but not the termination time T, the experiment can end much earlier than expected or last longer than expected. This is problematic in situations where experimenters expected or budgeted for a longer study. For example, in healthcare, participants may have been promised funding for a fixed duration, leading to wasted resources if the study ends prematurely. However, if the experiment lasts longer than expected, there may not be enough resources to complete the study.

In other cases, researchers may be specifically interested in late-life failures. For instance, when studying the durability of a manufactured product, it may be useful to know how likely it is to last well beyond its average lifespan. More details regarding the issues with Type-I and Type-II censoring are elaborated in Section 1.3 of Balakrishnan et al. [9].

With regards to Progressive Type-I and Type-II censoring, it is often difficult in practice to randomly withdraw additional live units at exactly the prescribed times (Type-I) or after each observed failure (Type-II). These challenges motivate the development of **hybrid censoring schemes**, which aim to balance the number of planned failures r and the termination time T while avoiding excessive complexity.

2.1.4 Hybrid Censoring Schemes

Two different types of hybrid censoring schemes are **Type-I Hybrid Censoring Scheme** and **Type-II Hybrid Censoring Scheme**. Typically, "Hybrid Censoring Scheme" is abbreviated to "HCS".

Suppose we fix both a termination time T and a pre-determined number of observed failures r. Let X_1, X_2, $\ldots$, X_n represent a random sample of lifetimes. We order the lifetimes as $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$. **Type-I Hybrid Censoring Scheme (HCS)** is when the experiment is terminated at $T_1 = \min\{X_{r:n}, T\}$ and is first introduced by Epstein [15]. One potential issue is that if T is small enough, it is possible that some estimators, such as the MLE, will not exist. This is because if T is sufficiently small, no failures may occur within the study period, leading to no data being collected, and therefore, inferences cannot be made.

Thus, other researchers may find the **Type-II Hybrid Censoring Scheme (HCS)** to be more appealing. In this scheme, lifetime experiments are terminated at $T_2 = \max\{X_{r:n}, T\}$, which provides more information than lifetime experiments under a Type-I HCS because it guarantees that there will be at least r failures. A limitation for Type-II HCS is that the experiment will need to run for longer compared to other censoring types, which can be seen as impractical depending on the time it takes for r observed failures to occur.

There are many other hybrid censoring schemes, such as Generalized Type-I HCS, Generalized Type-II HCS, and others. More details of alternative hybrid schemes can be found in Balakrishnan et al. [9].

2.1.5 Truncation

Truncated distributions are a type of conditional distribution that is derived from knowing that the variable has a smaller range than the initial distribution. For example, suppose we assume that the random sample $X_1, X_2, \ldots, X_n$ is exponentially distributed, but we know all of the lifetimes occur before time T. For reference, we say that this sample is **right-truncated** at T. Then, the range is $0 < X_i < T$ for $i \in \{1, 2, \ldots, n\}$. We need to adjust the probability density function to ensure it is valid. That is, letting $c \in \mathbb{R}$ represent the scalar that adjusts the probability density function accordingly:

$$1 = c \int_0^T \frac{1}{\theta} e^{-x/\theta} dx \quad \Rightarrow \quad c = (1 - e^{-T/\theta})^{-1}. \tag{2.1.12}$$

The truncation does not necessarily have to be between $0 < X_i < T$. Another form of truncation arises when we have a stricter lower bound than that of the original distribution. For example, consider the earlier sample $X_1, X_2, \ldots, X_n$ drawn from an exponential distribution, but this time all observed failures are known to exceed T. Consequently, the support becomes $T < X_i < \infty$ for $i \in \{1, 2, \ldots, n\}$, and such a sample is referred to as **left-truncated** at T.

Data can be doubly left and right-truncated. Hence, we highlight the general form of a truncated distribution. Suppose X has a probability density function $f_X(x)$ and a cumulative density function $F_X(x)$. Let $a, b \in \mathbb{R}$. If X follows a distribution restricted to the interval $[a, b]$, then the truncated distribution is defined as:

$$f_X(x \,|\, a \leq X \leq b) = \frac{f_X(x) \cdot \mathbf{1}_{[a,b]}(x)}{F_X(b) - F_X(a)}, \tag{2.1.13}$$

where $\mathbf{1}_{[a,b]}(x)$ represents the indicator function:

$$\mathbf{1}_{[a,b]}(x) = \begin{cases} 1 & \text{if } x \in [a, b], \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.14}$$

Truncations are relevant since they arise as a byproduct of censoring schemes. For example, under Type-I censoring, the experiment is forced to end at a fixed time T, which implies there are some right-truncated observations. Moreover, working with truncated distributions may be useful when dealing with estimators or mathematical expressions that include order statistics. For instance, it is useful to partition the ordered samples into two groups (e.g., one subject to right-truncation and the other to left-truncation).

When working with left-truncated samples from an exponential distribution, there is a convenient property: if the sample is truncated at some $T > 0$, then subtracting T from the truncated observations restores the original exponential distribution.

Lemma 2.1.3. *Let X be exponentially distributed with scale parameter θ, left-truncated at T. Then the shifted random variable $V = X - T$ also follows an exponential distribution with scale parameter θ.*

Proof. From Equation (2.1.13), the pdf of X is given by

$$f_X(x) = \frac{\theta^{-1}\exp(-x/\theta)}{1 - F_X(T)} = \frac{1}{\theta}\exp\left\{\frac{-x+T}{\theta}\right\}, \quad T < x < \infty. \tag{2.1.15}$$

Now, for $V = X - T$ we have

$$\begin{aligned} F_V(v) &= \mathbb{P}(X - T \le v) = F_X(v + T), \\ f_V(v) &= f_X(v + T) = \frac{1}{\theta}\exp\left\{\frac{-(v + T) + T}{\theta}\right\} \\ &= \frac{1}{\theta}\exp(-v/\theta)\,, \quad 0 < v < \infty, \end{aligned} \tag{2.1.16}$$

which is precisely the pdf of an exponential distribution with scale θ. □

Furthermore, when dealing with samples from a doubly truncated exponential distribution, it is often advantageous to transform the variable to obtain a distribution that is only right-truncated. Consider the following lemma as well.

Lemma 2.1.4. *Assume X is exponentially distributed with scale parameter θ and doubly truncated between T and T^* where $0 < T < T^*$. Then, $X - T$ is a right-truncated exponential distribution with scale parameter θ at $T^* - T$.*

Proof. From Equation (2.1.13), the pdf of X is given by

$$f_X(x) = \frac{\theta^{-1}\exp(-x/\theta)}{F_X(T^*) - F_X(T)} = \frac{\theta^{-1}\exp(-x/\theta)}{e^{-T/\theta} - e^{-T^*/\theta}}, \quad T < x < T^*. \tag{2.1.17}$$

Now, for $V = X - T$ we have

$$\begin{aligned} F_V(v) &= \mathbb{P}(X - T \leq v) = F_X(v + T), \\ f_V(v) &= f_X(v + T) \\ &= \frac{\theta^{-1}\exp(-(v+T)/\theta)}{e^{-T/\theta} - e^{-T^*/\theta}} \\ &= \frac{\theta^{-1}\exp(-v/\theta)\exp(-T/\theta)}{e^{-T/\theta} - e^{-T^*/\theta}} \frac{\exp(-T/\theta)^{-1}}{\exp(-T/\theta)^{-1}} \\ &= \frac{\theta^{-1}\exp(-v/\theta)}{1 - \exp(-(T^* + T)/\theta)} \end{aligned} \tag{2.1.18}$$

which follows the form of the pdf of an exponential distribution that is right-truncated at $T^* - T$ with scale θ. □

There is a useful theorem proved by Balakrishnan and Cohen [5] as Theorems (2.4.1) and (2.4.2) in their text, which shows that a sample of order statistics can be partitioned into two distinct groups: the first forming a right-truncated complete sample, and the second forming a left-truncated complete sample.

Theorem 2.1.2. *Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables from a population with cdf $F(x)$, and let $X_{1:n} \leq \cdots \leq X_{n:n}$ denote the corresponding order statistics. If the observed data are truncated on the right at T, and exactly m of the X_i's are less than or equal to T, then the first m order statistics $X_{1:n} \leq \cdots \leq X_{m:n}$ form a complete random sample of size m from the distribution $F(x)$ that is right-truncated at T. Furthermore, the remaining $n - m$ order statistics $X_{m+1:n} \leq \cdots \leq X_{n:n}$ form a complete random sample of size $n - m$ from the distribution $F(x)$ that is left-truncated at T. Moreover, these two sets of order statistics are independent.*

More details of truncation can again be found in Lawless' textbook [19].

2.2 Maximum Likelihood Estimation

This subchapter discusses likelihood functions, which are commonly used in statistical inference, together with their maximum likelihood estimators (MLEs). Since the form of the likelihood depends on the type of data, we also present the corresponding likelihoods under Type-I censoring, Type-II censoring, and the two hybrid schemes (Type-I and Type-II).

2.2.1 Likelihood Functions

Some methods in statistical inference rely on likelihood functions derived from observed data. Consider a random sample of lifetimes $X_1, X_2, \ldots, X_n$. These lifetimes are assumed to follow a distribution with pdf $f_X(x)$ and cdf $F_X(x)$ for $i = 1, 2, \ldots, n$. When we wish to emphasise the role of the parameters, we instead write $f_X(x; \boldsymbol{\theta})$ for a parameter vector $\boldsymbol{\theta}$. The likelihood function for complete data is then given by:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f_X(x_i; \boldsymbol{\theta}). \tag{2.2.1}$$

Because the likelihood depends on the observed data, its form differs under censoring schemes with incomplete observations, or with other missing-data mechanisms (e.g., data missing at random).

Suppose we have a random sample of lifetimes $X_1, X_2, \ldots, X_n$, with realisations $x_1, x_2, \ldots, x_n$. Each x_i represents either a fully observed lifetime or the time at which the observation was right-censored (i.e., the last known time before the experiment ended). Define a censoring indicator by:

$$\delta_i = \mathbf{1}_{x_i}(X_i) = \begin{cases} 1 & \text{if } X_i = x_i, \\ \\ 0 & \text{otherwise,} \end{cases} \quad \text{for } i = 1, 2, \ldots, n. \tag{2.2.2}$$

Thus, $\delta_i = 1$ indicates that the lifetime was observed in full, while $\delta_i = 0$ indicates censoring. Clearly, $\sum_{i=1}^{n} \delta_i$ gives the number of observed lifetimes.

The likelihood function can then be written as:

$$L(\theta) = \prod_{i=1}^{n} f_X(x_i)^{\delta_i} S_X(x_i)^{1-\delta_i}, \tag{2.2.3}$$

where $S_X(x_i) := \mathbb{P}(X \geq x_i)$ is the survival function. This form is intuitive; when $\delta_i = 1$, the observed lifetime contributes through the density $f_X(x_i)$. Meanwhile, if $\delta_i = 0$ then we only know that the lifetime exceeds x_i, which is captured by the survival probability $S_X(x_i)$. A more detailed discussion of likelihood functions for censored data can be found in Lawless' book [19].

Sometimes in practice, likelihood functions can be complex, and the closed-form expression for the MLE may not exist. In such cases, iterative procedures are employed. The Expectation-Maximisation (EM) algorithm is one widely used approach, and numerical methods such as Newton–Raphson can also be incorporated within it to carry out the maximisation step. A convenient feature of the exponential distribution is that the closed form for the likelihood functions and MLE can be easily derived under common right-censoring schemes, including the Type-II hybrid censoring scheme, so such numerical methods are not required for this thesis.

In the following subsections, we focus on the exponential distribution and explicitly derive the likelihood functions under four censoring schemes: Type-I censoring, Type-II censoring, Type-I hybrid censoring, and Type-II hybrid censoring.

2.2.2 Type-I Censoring

We now present a more formal definition of Type-I censoring. In this scheme, each lifetime observation has a pre-specified censoring time C_i. We use the term potential censoring time because if the event (failure) occurs before C_i, then the observation is not censored.

Assume now that $X_1, X_2, \ldots, X_n$ are i.i.d. and follow an exponential distribution with scale parameter θ, and let $x_1, x_2, \ldots, x_n$ denote the corresponding realisations. Additionally, let $r = \sum_{i=1}^{n} \delta_i = \sum_{i=1}^{n} \mathbf{1}_{x_i}(X_i)$ denote the number of complete (uncensored) observations. Then, the observed likelihood function is:

$$L(\theta) = \theta^{-r} \exp\left\{-\frac{1}{\theta} \sum_{i=1}^{n} \min\{x_i, C_i\}\right\}, \tag{2.2.4}$$

and the corresponding observed MLE is:

$$\hat{\theta}_{MLE} = \frac{r}{\sum_{i=1}^{n} \min\{x_i, C_i\}}. \tag{2.2.5}$$

Further details of this example can be found in Lawless' book [19].

2.2.3 Type-II Censoring

Consider a random sample $X_1, X_2, \ldots, X_n$ with realisations $x_1, x_2, \ldots, x_n$, each representing either a complete lifetime or a censoring time. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the corresponding order statistics with realisations $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$.

Under Type-II censoring, only the r smallest lifetimes are observed. Thus, $x_{1:n} \leq \cdots \leq x_{r:n}$ represent fully observed lifetimes, while $x_{r+1:n} \leq \cdots \leq x_{n:n}$ represent censored lifetimes, which are only known to exceed the censoring time $x_{r:n}$.

Assume that $X_1, \ldots, X_n$ are independent and exponentially distributed with scale parameter θ. Then, the observed likelihood function is:

$$L(\theta) = \theta^{-r} \exp\left\{-\frac{1}{\theta}\left[\sum_{i=1}^{r} x_{i:n} + (n-r)x_{r:n}\right]\right\}, \tag{2.2.6}$$

and accordingly, the observed MLE takes the form:

$$\hat{\theta}_{\text{MLE}} = \frac{r}{\sum_{i=1}^{r} x_{i:n} + (n-r)x_{r:n}}. \tag{2.2.7}$$

Again, comprehensive details are available in Lawless [19].

2.2.4 Type-I Hybrid Censoring

Consider a random sample $X_1, X_2, \ldots, X_n$. Under Type-I hybrid censoring, the lifetime experiment is terminated at $T_1 = \min\{X_{r:n}, T\}$ where $T \in (0, \infty)$ represents a censoring time, which is fixed in advance, and $r \in \mathbb{N}$ represents the number of failures we plan to observe.

Assume that D_1 corresponds to the number of failures before time T_1. Chen and

Bhattacharyya derived the likelihood function and the MLE for θ under the Type-I HCS censoring scheme, assuming the exponential distribution [10]:

$$L(\theta) = \begin{cases} \frac{n!}{(n-D_1)!}\theta^{-D_1} \exp\left\{-\frac{\sum_{i=1}^{D_1} X_i + (n-D_1)T_1}{\theta}\right\} & \text{if } D_1 \geq 1, \\ \exp\left\{-\frac{nT}{\theta}\right\} & \text{if } D_1 = 0, \end{cases} \tag{2.2.8}$$

and the corresponding MLE, which only exists when $D_1 > 0$, is:

$$\hat{\theta}_{MLE} = \begin{cases} \frac{1}{D}\left\{\sum_{i=1}^{D} X_i + (n-D)T\right\} & \text{if} \quad T < X_{r:n}, \\ \frac{1}{r}\left\{\sum_{i=1}^{r} X_i + (n-r)X_{r:n}\right\} & \text{if} \quad X_{r:n} \leq T. \end{cases} \tag{2.2.9}$$

Here, $D \in \{0, 1, 2, \ldots, n\}$ denotes the random number of failures that before time T, as illustrated in Figure 2.1.

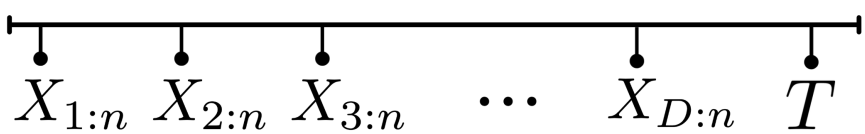

Figure 2.1: Timeline illustrating D, the number of failures occurring before time T.

One limitation of the Type-I HCS is that inferential results, such as the MLE, may not always exist. Notice that if the termination time T is small enough so that $D = 0$, then the MLE shown in Equation (2.2.9) will not exist. Hence, there is a need for an alternative.

2.2.5 Type-II Hybrid Censoring

To guarantee that the MLE will always exist, Childs et al. [11] introduced the **Type-II Hybrid Censoring Scheme**, where the experiment is terminated at $\max\{X_{r:n}, T\}$. Again, $X_{r:n}$ denotes the time of the r-th smallest failure among n observations, and T is a termination time set by the experimenter. This guarantees that there will be at least r observed failures for analysis. Figure 2.2 demonstrates the difference between Type-I and Type-II HCS.

Assuming that the lifetimes are exponentially distributed, the expression for the observed likelihood function for θ under the Type-II HCS censoring scheme is [11]:

$$L(\theta) = \begin{cases} \frac{n!}{(n-r)!}\frac{1}{\theta^r}\exp\left\{-\frac{1}{\theta}\sum_{i=1}^{r} x_i + (n-r)x_{r:n}\right\} & \text{if } D = 0, 1, \ldots, r-1, \\ \frac{n!}{(n-D)!}\frac{1}{\theta^D}\exp\left\{-\frac{1}{\theta}\sum_{i=1}^{D} x_i + (n-D)T\right\} & \text{if } D = r, r+1, \ldots, n, \end{cases} \tag{2.2.10}$$

and the corresponding observed MLE is:

$$\hat{\theta}_{MLE} = \begin{cases} \frac{1}{r}\left\{\sum_{i=1}^{r} x_i + (n-r)x_{r:n}\right\} & \text{if} \quad D = 0, 1, \ldots, r-1, \\ \frac{1}{D}\left\{\sum_{i=1}^{D} x_i + (n-D)T\right\} & \text{if} \quad D = r, r+1, \ldots, n. \end{cases} \tag{2.2.11}$$

2.3 Pitman Closeness Criterion

In this section, we present the Pitman closeness (PC) criterion, which serves as the primary tool in this thesis for comparing three different MLEs under the Type-II HCS. We also provide a motivation for using the PC criterion. We describe why

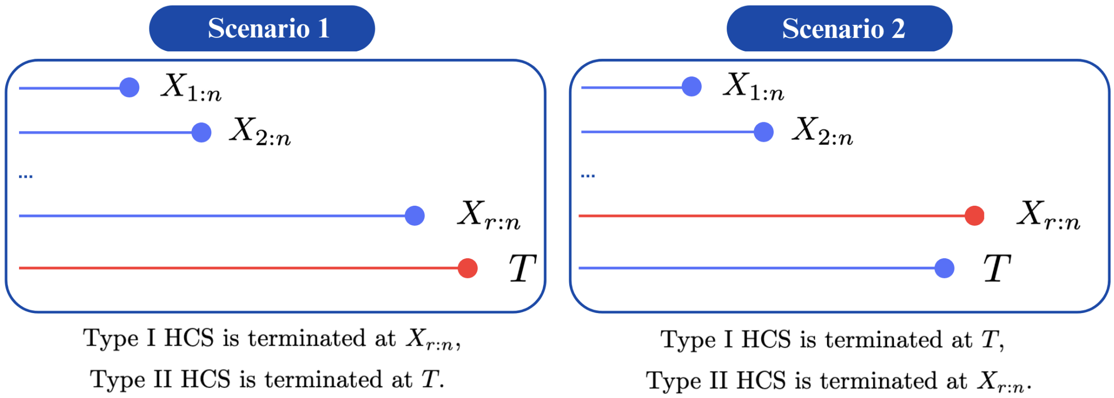

Figure 2.2: Illustration of the two different scenarios and the termination times for both Type-I HCS and Type-II HCS.

it offers a meaningful alternative to commonly used measures such as the mean squared error. Finally, we review the literature where PC has been applied to compare different estimators under various censoring schemes.

2.3.1 Definition

Suppose we want to compare two estimators, $\hat{\theta}_1$ and $\hat{\theta}_2$, of a common parameter θ. The **Pitman closeness criterion**, also known as the **Pitman measure of closeness**, is a method to compare two statistical estimators by computing the probability that the estimator $\hat{\theta}_1$ produces an estimate that is at least as close to the true value θ as the estimate given by the second estimator $\hat{\theta}_2$. To clarify a common misconception, PC does not evaluate the magnitude of how close an estimator is to the parameter.

In 1936, Pearson questioned what it means for an estimator to be "better" among two competing estimators [20]. In response, the PC criterion was introduced in the same year [21]. Ideally, a "better" estimator is the one that is more likely to produce estimates closer to the true parameter value. Later, a textbook was dedicated to

Pitman closeness [18]. Specifically, let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two competing estimators of a parameter θ. Then, the PC of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is:

$$\pi_{\hat{\theta}_1,\hat{\theta}_2}(\theta) := \mathbb{P}(|\hat{\theta}_1 - \theta| \leq |\hat{\theta}_2 - \theta|). \tag{2.3.1}$$

The PC criterion is as follows: if $\pi_{\hat{\theta}_1,\hat{\theta}_2}(\theta) > \frac{1}{2}$ then $\hat{\theta}_1$ is said to be **Pitman closer** to θ, implying it is a more desirable estimator.

2.3.2 Motivation and Rationale

There are many different methods for comparing estimators; the most commonly taught method is minimising the mean square error (MSE). Let $\hat{\theta}$ be an estimator with respect to an unknown parameter θ. Then, the MSE is defined as:

$$MSE(\hat{\theta}) := \mathbb{E}\left[(\hat{\theta} - \theta)^2\right]. \tag{2.3.2}$$

It is common to use the formula pertaining to the variance-bias decomposition:

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + Bias(\hat{\theta})^2. \tag{2.3.3}$$

Karlin [16] noted advantages of using MSE when comparing estimators: for an unbiased estimator $\hat{\theta}$ of θ, the MSE equals its variance, and squared error emphasises large deviations, so a smaller MSE implies more consistent accuracy. However, Rao [22] argued that MSE does not indicate how often an estimator lies close to the true parameter.

Focusing solely on MSE encourages reducing the variance of unbiased estimators, hence the origin of solving for the unique minimum variance unbiased estimators (UMVUE). However, many useful estimators, such as the MLE, may be biased. Rao [22] therefore proposed PC as an alternative, since it directly measures the probability that one estimator is closer (in absolute error) to the true parameter than another. He demonstrated cases where an unbiased estimator with smaller variance does not yield a Pitman closer estimator compared to one with higher variance, illustrating a need for alternative means for comparing estimators.

The PC also has an intuitive appeal in applied settings. For example, Keating and Mason [17] noted that in elections, voters typically support the candidate whose position is likely the closest to their own, not the one minimising squared differences. Similarly, customers choose the convenience store that will probably take the shortest amount of time to travel, not the one that minimises a weighted squared distance. In practice, decisions are often made based on closeness rather than MSE.

The textbook by Keating et al. [18] provides additional cases to illustrate why the PC criterion can be regarded as more intuitive than alternative approaches under different settings. Furthermore, recent literature has included examples of computing the PC probabilities of different estimators (such as the MLE, median predictor, etc.) when assuming the data follows an exponential distribution [1, 2].

2.3.3 Applications in the Literature

Several studies have applied the PC criterion to compare different estimators under different censoring schemes. For instance, under Type-I censoring of exponential lifetime data, Balakrishnan et al. [7] considered two MLEs of the scale parameter θ corresponding to different termination times, T and T^*, where $T < T^*$. They found that the MLE based on the longer termination time, T^*, was always Pitman closer to the true parameter than the one based on T for the cases considered.

Furthermore, under Type-II censoring of exponential lifetime data, Balakrishnan et al. [8] further showed that, in the cases they considered, the best linear unbiased estimator (BLUE) of the scale parameter θ is always Pitman closer to θ than the best linear invariant estimator (BLIE) of θ.

Recently, the PC probabilities for the MLEs of the scale parameter θ under Type-I HCS, assuming an exponential distribution, has been computed [13]. There were two comparisons: the first was between two estimators based on Type-I HCS with differing termination times $\min\{X_{r:n}, T\}$ and $\min\{X_{r:n}, T^*\}$ where $T < T^*$. In the specific cases examined by the author, it appeared that the estimator with the longer termination time, T^*, usually produced an estimator that is Pitman closer to θ than the shorter one, with rare exceptions.

In the second comparison, the number of observed failures was varied while keeping the termination time fixed, comparing $\min\{X_{r:n}, T\}$ and $\min\{X_{s:n}, T\}$ with $r < s$. The author found that increasing the number of failures observed before stopping almost always led to an estimator that was Pitman closer to θ. Both of these comparisons align with the intuition that a longer experiment or more data produced estimates that are Pitman closer to the true parameter.

We now extend these comparisons by applying the PC criterion to MLEs of the scale parameter θ under Type-II hybrid censoring schemes (HCS) from an exponential distribution, a case that has not been considered. In our first comparison, we consider the MLEs $\hat{\theta}_1$ and $\hat{\theta}_2$, which are based on the termination times $\max\{X_{r:n}, T\}$ and $\max\{X_{s:n}, T\}$, respectively, where $r < s$. In the second comparison, we contrast $\hat{\theta}_1$ with $\hat{\theta}_3$, where $\hat{\theta}_3$ is based on the termination time $\max\{X_{r:n}, T^*\}$ with $T < T^*$.

Chapter 3

Comparison of $\hat{\theta}_1$ and $\hat{\theta}_2$

This chapter computes the Pitman closeness (PC) of two estimators based on Type-II HCS experiments. We investigate which estimator is Pitman closer to the scale parameter θ of the exponential distribution once we increase the number of observed failures.

3.1 Estimators and Case Breakdown

Assume we have a lifetime experiment with a random sample of lifetimes X_1, X_2, $\ldots$, X_n which follows an exponential distribution. That is, let $X_i \sim \mathit{Exponential}(\theta)$ for $i \in \{1, 2, \ldots, n\}$. Then let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the corresponding order statistics. Due to the scale-invariant property under reparameterisation of the exponential distribution as described in Equation (2.1.7), we assume a rate parameter of $\theta = 1$ without loss of generality.

Suppose we have a pre-determined termination time T. Let D be a random variable that represents the number of failures during the interval $(0, T]$. That is, $D = \sum_{i=1}^{n} \mathbf{1}_{X_{i:n}}(0,T]$. Since we have a summation of the complete sample, the order does not matter due to the commutative property of addition, and so $D = \sum_{i=1}^{n} \mathbf{1}_{X_i}(0,T]$.

Clearly, $\mathbf{1}_{X_i}(0,T]$ is Bernoulli distributed where $\mathbb{P}(X_i \le T) = 1 - e^{-T}$. Since the sum of independent and identically Bernoulli random variables is binomial, we have that $D \sim Binomial(n, 1 - e^{-T})$.

Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be the respective MLEs for θ under the termination times $\max\{X_{r:n}, T\}$ and $\max\{X_{s:n}, T\}$, where $r < s$. Borrowing the results from Childs et al. [11], these respective estimators are of the form:

$$\hat{\theta}_1 = \begin{cases} \frac{1}{r}\left\{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right\} & D = 0, 1, \ldots, r-1, \\ \frac{1}{D}\left\{\sum_{i=1}^{D} X_{i:n} + (n-D)T\right\} & D = r, r+1, \ldots, n, \end{cases} \tag{3.1.1}$$

and

$$\hat{\theta}_2 = \begin{cases} \frac{1}{s}\left\{\sum_{i=1}^{s} X_{i:n} + (n-s)X_{s:n}\right\} & D = 0, 1, \ldots, s-1, \\ \frac{1}{D}\left\{\sum_{i=1}^{D} X_{i:n} + (n-D)T\right\} & D = s, s+1, \ldots, n. \end{cases} \tag{3.1.2}$$

Since the expressions for $\hat{\theta}_1$ and $\hat{\theta}_2$ both depend on the value of D, we need to consider the following cases:

Case 1: $D \in \{0, 1, \ldots, r-1\} \iff T < X_{r:n} < X_{s:n}$,

Case 2: $D \in \{r, r+1, \ldots, s-1\} \iff X_{r:n} \le T < X_{s:n}$,

Case 3: $D \in \{s, \ldots, n\} \iff X_{r:n} < X_{s:n} \le T$.

To account for these separate cases, we will condition on them accordingly based on the value of D. For case 1, we focus on the estimators:

$$\hat{\theta}_{111} := \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right), \quad \hat{\theta}_{112} := \frac{1}{s}\left(\sum_{i=1}^{s} X_{i:n} + (n-s)X_{s:n}\right). \tag{3.1.3}$$

Under case 2, the estimators take the form:

$$\hat{\theta}_{121} := \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right), \quad \hat{\theta}_{122} := \frac{1}{s}\left(\sum_{i=1}^{s} X_{i:n} + (n-s)X_{s:n}\right). \tag{3.1.4}$$

Finally, when considering case 3, the estimators are expressed as:

$$\hat{\theta}_{131} := \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right), \quad \hat{\theta}_{132} := \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right). \tag{3.1.5}$$

For notational convenience, let $\mathbb{P}_D(\cdot) := \mathbb{P}(\cdot|D = d)$. We can then derive the PC probability using the following decomposition:

$$\begin{aligned}\mathbb{P}(|\hat{\theta}_2 - 1| \le |\hat{\theta}_1 - 1|) = &\sum_{d=0}^{r-1} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{112} - 1| \le |\hat{\theta}_{111} - 1|)\\ &+ \sum_{d=r}^{s-1} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{122} - 1| \le |\hat{\theta}_{121} - 1|)\\ &+ \sum_{d=s}^{n} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{132} - 1| \le |\hat{\theta}_{131} - 1|).\end{aligned} \tag{3.1.6}$$

Since $D \sim Binomial(n, 1 - e^{-T})$, we can easily derive:

$$\mathbb{P}(D = d) = \binom{n}{d}(1 - e^{-T})^d(e^{-T})^{n-d}. \tag{3.1.7}$$

For the next subchapters, we will focus on computing the following cases:

Case 1: $\sum_{d=0}^{r-1} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{112} - 1| \leq |\hat{\theta}_{111} - 1|)$,

Case 2: $\sum_{d=r}^{s-1} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{122} - 1| \leq |\hat{\theta}_{121} - 1|)$,

Case 3: $\sum_{d=s}^{n} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{132} - 1| \leq |\hat{\theta}_{131} - 1|)$.

3.2 Case 1

As a reminder, the estimators we want to compare depend specifically on the values of D. Hence, we will condition on $D = d$ when computing the PC between $\hat{\theta}_1$ and $\hat{\theta}_2$.

For our first case, we consider the values where $D \in \{0, 1, \ldots, r-1\}$. These values of D correspond to these two estimators:

$$\hat{\theta}_{111} := \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right), \quad \hat{\theta}_{112} := \frac{1}{s}\left(\sum_{i=1}^{s} X_{i:n} + (n-s)X_{s:n}\right). \tag{3.2.1}$$

We will compute:

$$\sum_{d=0}^{r-1} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{112} - 1| \leq |\hat{\theta}_{111} - 1|). \tag{3.2.2}$$

To start, we consider the conditional PC probability:

$$\begin{aligned}
\mathbb{P}_D(|\hat{\theta}_{112} - 1| \leq |\hat{\theta}_{111} - 1|) &= \mathbb{P}_D((\hat{\theta}_{112} - \hat{\theta}_{111})(\hat{\theta}_{112} + \hat{\theta}_1 - 2) \leq 0) \\
&= \mathbb{P}_D(2 - \hat{\theta}_{111} \leq \hat{\theta}_{112} \leq \hat{\theta}_{111}) + \mathbb{P}_D(\hat{\theta}_{111} \leq \hat{\theta}_{112} \leq 2 - \hat{\theta}_{111}) \\
&= \pi_{111} + \pi_{112}, \quad \text{say,} \qquad (3.2.3)
\end{aligned}$$

where π_{111} and π_{112} are used to simplify the expression. First, we will rearrange the inequality within π_{111}.

$$
\begin{aligned}
\pi_{111} &= \mathbb{P}_D(2-\hat{\theta}_{111} \leq \hat{\theta}_{112} \leq \hat{\theta}_{111}) \\
&= \mathbb{P}_D\Bigg\{2 - \frac{1}{r}\Bigg(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Bigg) \leq \frac{1}{s}\Bigg(\sum_{i=1}^{s} X_{i:n} + (n-s)X_{s:n}\Bigg) \\
&\leq \frac{1}{r}\Bigg(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Bigg)\Bigg\} \\
&= \mathbb{P}_D\Bigg\{2 - \frac{1}{r}\Bigg(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Bigg) - \frac{1}{s}\Bigg(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Bigg) \\
&\leq \frac{1}{s}\Bigg(\sum_{i=1}^{s} X_{i:n} + (n-s)X_{s:n}\Bigg) - \frac{1}{s}\Bigg(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Bigg) \\
&\leq \frac{1}{r}\Bigg(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Bigg) - \frac{1}{s}\Bigg(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Bigg)\Bigg\} \\
&= \mathbb{P}_D\Bigg\{2 - \Big(\frac{1}{r} + \frac{1}{s}\Big)\sum_{i=1}^{r} X_{i:n} - \Big(\frac{n-r}{r} + \frac{n-r}{s}\Big)X_{r:n} \\
&\leq \frac{1}{s}\Bigg(\sum_{i=r+1}^{s} X_{i:n} + (n-s)X_{s:n} - (n-s+s-r)X_{r:n}\Bigg) \\
&\leq \Big(\frac{1}{r} - \frac{1}{s}\Big)\sum_{i=1}^{r} X_{i:n} + \Big(\frac{n-r}{r} - \frac{n-r}{s}\Big)X_{r:n}\Bigg\} \\
&= \mathbb{P}_D\Bigg\{s\Bigg[2 - \Big(\frac{1}{r} + \frac{1}{s}\Big)\sum_{i=1}^{r} X_{i:n} - \Big(\frac{n-r}{r} + \frac{n-r}{s}\Big)X_{r:n}\Bigg] \\
&\leq \sum_{i=r+1}^{s} (X_{i:n} - X_{r:n}) + (n-s)(X_{s:n} - X_{r:n}) \\
&\leq s\Bigg[\Big(\frac{1}{r} - \frac{1}{s}\Big)\sum_{i=1}^{r} X_{i:n} + \Big(\frac{n-r}{r} - \frac{n-r}{s}\Big)X_{r:n}\Bigg]\Bigg\}.
\end{aligned} \tag{3.2.4}
$$

Using similar steps, we obtain the following for π_{112}:

$$\pi_{112} = \mathbb{P}_D(\hat{\theta}_{111} \leq \hat{\theta}_{112} \leq 2-\hat{\theta}_{111})$$

$$
\begin{aligned}
&= \mathbb{P}_D\Bigg\{ s\Bigg[\left(\frac{1}{r}-\frac{1}{s}\right)\sum_{i=1}^{r-1} X_{i:n} + \left(\frac{n-r}{r}-\frac{n-r}{s}+\frac{1}{r}-\frac{1}{s}\right)X_{r:n}\Bigg] \\
&\leq \sum_{i=r+1}^{s}(X_{i:n}-X_{r:n}) + (n-s)(X_{s:n}-X_{r:n}) \\
&\leq s\Bigg[2-\left(\frac{1}{r}+\frac{1}{s}\right)\sum_{i=1}^{r-1} X_{i:n} - \left(\frac{n-r}{r}+\frac{n-r}{s}+\frac{1}{r}+\frac{1}{s}\right)X_{r:n}\Bigg]\Bigg\}. \qquad (3.2.5)
\end{aligned}
$$

Now, we will derive the pdf of the middle quantity of the inequality in the conditional PC probability shown in Equation (3.2.5). That is, we define:

$$
B_{11} := \sum_{i=r+1}^{s}(X_{i:n}-X_{r:n}) + (n-s)(X_{s:n}-X_{r:n}). \qquad (3.2.6)
$$

Conditioning on $X_{r:n} = x_r$, using Theorem (2.1.2), we claim that $X_{r+1:n} \leq X_{r+2:n} \leq \cdots \leq X_{s:n}$ are the order statistics from a complete sample of size $(s-r)$ from the exponential distribution truncated on the left at x_r. By Lemma (2.1.3), subtracting x_r from these truncated variables will produce random variables that are exponentially distributed. Hence, re-writing B_{11} provides us with:

$$
\begin{aligned}
B_{11} &= \sum_{i=r+1}^{s}(X_{i:n}-x_r) + (n-s)(X_{s:n}-x_r) \\
&= \sum_{i=1}^{s-r} Y_{i:n-r} + (n-s)Y_{s-r:n-r} \\
&= \sum_{i=1}^{s-r-1} Y_{i:n-r} + (n-s+1)Y_{s-r:n-r}, \qquad (3.2.7)
\end{aligned}
$$

where the random variables $Y_{1:n-r} \leq Y_{2:n-r} \leq \cdots \leq Y_{s-r:n-r}$ denote the first $(s-r)$ order statistics from a sample of size $(n-r)$ drawn from a standard exponential distribution. To find the distribution of B_{11}, we realise that B_{11} is the same form as the sum of spacings as shown in the form of Theorem (2.1.1) where we denote $Y_{0:n-r} = 0$. In particular,

$$
\begin{aligned}
&\sum_{i=1}^{s-r} ((n-r) - i + 1)(Y_{i:n-r} - Y_{i-1:n-r}) \\
&= (n-r)Y_{1:n-r} + (n-r-1)(Y_{2:n-r} - Y_{1:n-r}) + (n-r-2)(Y_{3:n-r} + Y_{2:n-r}) \\
&\quad + \cdots + \underbrace{(n-r-(s-r-1))}_{(n-s+1)}(Y_{s-r:n-r} - Y_{s-r-1:n-r}) \\
&= [(n-r) - (n-r-1)]Y_{1:n-r} + [(n-r-1) - (n-r-2)]Y_{2:n-r} \\
&\quad + \cdots + [(n-s+2) - (n-s+1)]Y_{s-r-1:n-r} + [n-s+1]Y_{s-r:n-r} \\
&= Y_{1:n-r} + Y_{2:n-r} + \cdots + Y_{s-r-1:n-r} + (n-s+1)Y_{s-r:n-r} \\
&= \sum_{i=1}^{s-r-1} Y_{i:n-r} + (n-s+1)Y_{s-r:n-r} \\
&= B_{11}.
\end{aligned} \tag{3.2.8}
$$

Thus, we can use Theorem (2.1.1) and Lemma (2.1.1) to claim that:

$$
B_{11} = \sum_{i=1}^{s-r} ((n-r) - i + 1)(Y_{i:n-r} - Y_{i-1:n-r}) = \sum_{i=1}^{s-r} Z_i \sim Gamma(s-r, 1). \tag{3.2.9}
$$

Now we are interested in observing the bounds of the inequality shown in Equations (3.2.4) and (3.2.5). Notice that the bounds depend on both $X_{r:n}$ and $\sum_{i=1}^{r-1} X_{i:n}$. Given that we are conditioning on $X_{r:n} = x_r$, Theorem (2.1.2) allows us to claim that the random variables $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{r-1:n}$ are distributed as the order statistics of a complete sample of size $(r-1)$ from an exponential distribution truncated on the right of x_r.

Hence, we will now try to find the distribution of $A_{11} \coloneqq \sum_{i=1}^{r-1} X_{i:n}$ given $X_{r:n} = x_r$. Since there will be many instances in which we will need to find the distribution of the sum of order statistics that are right-truncated, we will instead state a general proposition and prove it.

Proposition 3.2.1. ***Sum of Right-Truncated Ordered Exponential.*** *Let $X_{1:n} \leq \cdots \leq X_{n:n}$ be the order statistics obtained from a sample of size n from the Exponential(1) distribution. If the obtained data are truncated on the right at L, and exactly m of the X_i's are less than or equal to L, then, the first m order statistics $X_{1:n} \leq \cdots \leq X_{m:n}$ form a complete random sample of size m from the right-truncated exponential distribution at L. Additionally, the pdf of $A' = \sum_{i=1}^{m} X_{i:n}$ is:*

$$f_{A'}(a) = \frac{1}{(1-e^{-L})^m} \sum_{i=0}^{m} \binom{m}{i} (-1)^i e^{-Li} \frac{1}{\Gamma(m)} (a - Li)^{m-1} e^{-(a-Li)} I(a > Li), \qquad (3.2.10)$$

where $0 \leq A' \leq mL$. Furthermore, the last $n-m$ order statistics $X_{m+1:n} \leq \cdots \leq X_{n:n}$ form a complete random sample of size $n-m$ from the left-truncated exponential distribution at L. Moreover, the two sets of order statistics are independent.

Proof. We focus on deriving the pdf of A' as Theorem (2.1.2) covers the rest of the statement. Let X' represent a single observation that is exponentially distributed, which is right-truncated at L. Then, the density function of X' is:

$$f_{X'}(x) = \frac{e^{-x}}{1-e^{-L}}, \quad 0 \leq x \leq L. \qquad (3.2.11)$$

Its moment generating function is:

$$M_{X'}(t) = \mathbb{E}\left[e^{X't}\right]$$

$$
\begin{aligned}
&= \int_0^L e^{xt} \frac{e^{-x}}{(1-e^{-L})} dx \\
&= \frac{1}{(1-e^{-L})} \left[\frac{-e^{-x(1-t)}}{(1-t)} \Big|_0^L \right] \\
&= \frac{1}{(1-e^{-L})} \frac{1}{(1-t)} [1 - e^{-L(1-t)}].
\end{aligned} \tag{3.2.12}
$$

Thus, we can compute the moment generating function for A':

$$
\begin{aligned}
M_{A'} &= \left(\mathbb{E}\left[e^{X't} \right] \right)^m \\
&= \frac{1}{(1-e^{-L})^m} \frac{1}{(1-t)^m} [1 - e^{-L(1-t)}]^m \\
&= \frac{1}{(1-e^{-L})^m} \frac{1}{(1-t)^m} \sum_{i=0}^{m} \binom{m}{i} (-1)^i e^{-Li} e^{Lti} \\
&= \frac{1}{(1-e^{-L})^m} \sum_{i=0}^{m} \binom{m}{i} (-1)^i e^{-Li} \frac{e^{Lti}}{(1-t)^m}.
\end{aligned} \tag{3.2.13}
$$

The first line arises from having a complete sample of size m, and binomial theorem is used for the third line. To derive the pdf, we will work backwards by referring to common moment generating functions.

Suppose $W \sim Gamma(m, 1)$, and $i \in \mathbb{N}$. Then, the moment generating function of $Li + W$ is:

$$
M_{Li+W}(t) = e^{Lit} \mathbb{E}[e^{Wt}] = \frac{e^{Lit}}{(1-t)^m}. \tag{3.2.14}
$$

This allows us to find the cdf and pdf:

$$
\begin{aligned}
F_{Li+W}(w) &= \mathbb{P}(Li + W \leq w) = \mathbb{P}(W \leq w - Li) = F_W(w - Li) \\
f_{Li+W}(w) &= f_W(w - Li) = \frac{1}{\Gamma(m)} (w - Li)^{m-1} e^{-(w-Li)} I(w > Li).
\end{aligned} \tag{3.2.15}
$$

This is used as inspiration to propose a pdf for A':

$$f_{A'}(a) = \frac{1}{(1-e^{-L})^m} \sum_{i=0}^{m} \binom{m}{i} (-1)^i e^{-Li} \frac{1}{\Gamma(m)} (a - Li)^{m-1} e^{-(a-Li)} I(a > Li). \tag{3.2.16}$$

It is not difficult to show that computing the moment generating function using the proposed pdf, Equation (3.2.16), will have the same result as Equation (3.2.13). Hence, due to the uniqueness property of moment generating functions, the proposed distribution is valid. The next thing we are interested in is finding the bounds of A'.

Since A' is the sum of exponentially distributed random variables whose support is positive, $A' \geq 0$. Furthermore, since the observations are right-truncated at L we have that $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{m:n} \leq L$. The maximum case occurs when we have equality. That is, $X_{1:n} = X_{2:n} = \cdots = X_{m:n} = L$, so $A' \leq mL$. □

Using Proposition (3.2.1), we know that conditioning on $X_{r:n} = x_r$ tells us that $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{r-1:n} \leq x_r$ and therefore if we set $m = r - 1$ and $L = x_r$ we have that the pdf of A_{11} is:

$$f_{A_{11}|X_{r:n}=x_r}(a) = \frac{1}{(1-e^{-x_r})^{r-1}} \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i e^{-x_r i} \frac{1}{\Gamma(r-1)} (a - x_r i)^{(r-1)-1} e^{-(a-x_r i)} I(a > x_r i). \tag{3.2.17}$$

Notably, the bounds shown within the probabilities defined by π_{111} and π_{112} (see Equations (3.2.4) and (3.2.5), respectively) suggest different values for A_{11}. For notational convenience, let:

$$L_{11}(A_{11}, x_r) := s\left[2 - \left(\frac{1}{r} + \frac{1}{s}\right) A_{11} - \left(\frac{n-r}{r} + \frac{n-r}{s} + \frac{1}{r} + \frac{1}{s}\right) x_r\right], \tag{3.2.18}$$

$$U_{11}(A_{11}, x_r) := s\left[\left(\frac{1}{r} - \frac{1}{s}\right)A_{11} + \left(\frac{n-r}{r} - \frac{n-r}{s} + \frac{1}{r} - \frac{1}{s}\right)x_r\right]. \tag{3.2.19}$$

First, we consider the case corresponding to π_{111}.

$$\begin{aligned}
&L_{11}(A_{11}, x_r) \leq U_{11}(A_{11}, x_r) \\
&\Leftrightarrow s\left[2 - \left(\frac{1}{r} + \frac{1}{s}\right)A_{11} - \left(\frac{n-r}{r} + \frac{n-r}{s} + \frac{1}{r} + \frac{1}{s}\right)x_r\right] \\
&\leq s\left[\left(\frac{1}{r} - \frac{1}{s}\right)A_{11} + \left(\frac{n-r}{r} - \frac{n-r}{s} + \frac{1}{r} - \frac{1}{s}\right)x_r\right] \\
&\Leftrightarrow 2 - \left(\frac{1}{r} + \frac{1}{s}\right)A_{11} - \left(\frac{1}{r} - \frac{1}{s}\right)A_{11} \\
&\leq x_r\left(\frac{n-r}{r} \cancel{- \frac{n-r}{s}} + \frac{1}{r} \cancel{- \frac{1}{s}} + \frac{n-r}{r} \cancel{+ \frac{n-r}{s}} + \frac{1}{r} \cancel{+ \frac{1}{s}}\right) \\
&\Leftrightarrow 2 - \frac{2A_{11}}{r} \leq 2x_r\frac{n-r+1}{r} \\
&\Leftrightarrow r - x_r(n-r+1) \leq A_{11}.
\end{aligned} \tag{3.2.20}$$

Similarly, for the case corresponding to π_{112}, $U_{11}(A_{11}, x_r) \leq L_{11}(A_{11}, x_r) \quad \Leftrightarrow \quad A_{11} \leq r - x_r(n-r+1)$.

Next, we will propose a pdf for $X_{r:n}$ conditional on $D = d$. Since we have that $D \in \{0, 1, \ldots, r-1\}$ it is known that $X_{r:n} > T$. Let X''_{11} denote a random variable following a standard exponential distribution left-truncated at T. Then, its pdf and cdf are given by:

$$f_{X''_{11}}(x) = \frac{e^{-x}}{e^{-T}}, \text{ and } F_{X''_{11}}(x) = 1 - \frac{e^{-x}}{e^{-T}}, \tag{3.2.21}$$

respectively. Given that we are conditioning on $D = d$, $X_{r:n}$ is the $(r-d)$th order statistic from $(n-d)$ observations left-truncated at T. Thus, we can use Lemma (2.1.2) to derive the density function of $X_{r:n}$ conditional on $D = d$:

$$
\begin{aligned}
f_{X_{r:n}|D=d}(v) &= \frac{(n-d)!}{((r-d)-1)!((n-d)-(r-d))!}\{F_{X_r}(v)\}^{(r-d)-1}\{1-F_{X_r}(v)\}^{(n-d)-(r-d)} f_{X_r}(v) \\
&= \frac{(n-d)!}{(r-d-1)!(n-r)!}\left(1-\frac{e^{-v}}{e^{-T}}\right)^{r-d-1}\left(\frac{e^{-v}}{e^{-T}}\right)^{n-r}\left(\frac{e^{-v}}{e^{-T}}\right) \\
&= \frac{(n-d)!}{(r-d-1)!(n-r)!}\left(1-\frac{e^{-v}}{e^{-T}}\right)^{r-d-1}\left(\frac{e^{-v}}{e^{-T}}\right)^{n-r+1}, \quad T < v < \infty. \qquad (3.2.22)
\end{aligned}
$$

Now, we can derive π_{111} and π_{112} using the total law of probability:

$$
\begin{aligned}
\pi_{111} &= \mathbb{P}_D\Bigg\{s\Bigg[2-\left(\frac{1}{r}+\frac{1}{s}\right)\sum_{i=1}^{r} X_{i:n} - \left(\frac{n-r}{r}+\frac{n-r}{s}\right)X_{r:n}\Bigg] \\
&\leq \left(\sum_{i=r+1}^{s}(X_{i:n}-X_{r:n}) + (n-s)(X_{s:n}-X_{r:n})\right) \\
&\leq s\Bigg[\left(\frac{1}{r}-\frac{1}{s}\right)\sum_{i=1}^{r} X_{i:n} + \left(\frac{n-r}{r}-\frac{n-r}{s}\right)X_{r:n}\Bigg]\Bigg\} \\
&= \int_T^{\infty}\int_{r-x_r(n-r+1)}^{(r-1)x_r} \mathbb{P}_D\left\{L_{11}(a,x_r) \leq B_{11} \leq U_{11}(a,x_r)\,|\,X_{r:n}=x_r, A_{11}=a\right\} \\
&\quad \times f_{X_{r:n},A_{11}|D=d}(x_r,a)\,da\,dx_r \\
&= \int_T^{\infty}\int_{r-x_r(n-r+1)}^{(r-1)x_r} \Big[F_{B_{11}}(U_{11}(a,x_r)) - F_{B_{11}}(L_{11}(a,x_r))\Big] \\
&\quad \times f_{X_{r:n},A_{11}|D=d}(x_r,a)\,da\,dx_r \\
&= \int_T^{\infty}\int_{r-x_r(n-r+1)}^{(r-1)x_r} \Big[F_{B_{11}}(U_{11}(a,x_r)) - F_{B_{11}}(L_{11}(a,x_r))\Big] \\
&\quad \times f_{A_{11}|X_{r:n}=x_r,D=d}(a)\, f_{X_{r:n}|D=d}(x_r)\,da\,dx_r, \qquad (3.2.23)
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
\pi_{112} = \mathbb{P}_D\Bigg\{ & s\Bigg[\left(\frac{1}{r}-\frac{1}{s}\right)\sum_{i=1}^{r-1} X_{i:n} + \left(\frac{n-r}{r}-\frac{n-r}{s}+\frac{1}{r}-\frac{1}{s}\right)X_{r:n}\Bigg] \\
& \le \left(\sum_{i=r+1}^{s}(X_{i:n}-X_{r:n}) + (n-s)(X_{s:n}-X_{r:n})\right) \\
& \le s\Bigg[2-\left(\frac{1}{r}+\frac{1}{s}\right)\sum_{i=1}^{r-1} X_{i:n} - \left(\frac{n-r}{r}+\frac{n-r}{s}+\frac{1}{r}+\frac{1}{s}\right)X_{r:n}\Bigg]\Bigg\} \\
& = \int_T^{\infty}\int_0^{r-x_r(n-r+1)} \mathbb{P}_D\left\{U_{11}(a,x_r) \le B_{11} \le L_{11}(a,x_r)\,|\,X_{r:n}=x_r, A_{11}=a\right\} \\
& \quad \times f_{X_{r:n},A_{11}|D=d}(x_r,a)\,da\,dx_r \\
& = \int_T^{\infty}\int_0^{r-x_r(n-r+1)} \left[F_{B_{11}}(L_{11}(a,x_r)) - F_{B_{11}}(U_{11}(a,x_r))\right] \\
& \quad \times f_{X_{r:n},A_{11}|D=d}(x_r,a)\,da\,dx_r \\
& = \int_T^{\infty}\int_0^{r-x_r(n-r+1)} \left[F_{B_{11}}(L_{11}(a,x_r)) - F_{B_{11}}(U_{11}(a,x_r))\right] \\
& \quad \times f_{A_{11}|X_{r:n}=x_r,D=d}(a)\,f_{X_{r:n}|D=d}(x_r)\,da\,dx_r.
\end{aligned}
\tag{3.2.24}
$$

Therefore, the PC probability corresponding to the first case is:

$$
\sum_{d=0}^{r-1}\mathbb{P}(D=d)\mathbb{P}_D(|\hat{\theta}_{112}-1| \le |\hat{\theta}_{111}-1|) = \sum_{d=0}^{r-1}\mathbb{P}(D=d)[\pi_{111}+\pi_{112}]. \tag{3.2.25}
$$

3.3 Case 2

For case 2, the objective is to compare the following two estimators, which arise when $D \in \{r, r+1, \ldots, s-1\}$:

$$\hat{\theta}_{121} := \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right), \quad \hat{\theta}_{122} := \frac{1}{s}\left(\sum_{i=1}^{s} X_{i:n} + (n-s)X_{s:n}\right). \tag{3.3.1}$$

Again, we attempt to derive the conditional PC probability:

$$\begin{aligned}\mathbb{P}_D(|\hat{\theta}_{122} - 1| \leq |\hat{\theta}_{121} - 1|) &= \mathbb{P}_D((\hat{\theta}_{122} - \hat{\theta}_{121})(\hat{\theta}_{122} + \hat{\theta}_{121} - 2) \leq 0) \\ &= \mathbb{P}_D(2 - \hat{\theta}_{121} \leq \hat{\theta}_{122} \leq \hat{\theta}_{121}) + \mathbb{P}_D(\hat{\theta}_{121} \leq \hat{\theta}_{122} \leq 2 - \hat{\theta}_{121}) \\ &= \pi_{121} + \pi_{122}, \text{ say,} \end{aligned} \tag{3.3.2}$$

where π_{121} and π_{122} are named for convenience. To continue our derivation, we must find a way to simplify $\hat{\theta}_{122}$. Note that we can re-arrange $\hat{\theta}_{122}$ to have the form:

$$\hat{\theta}_{122} = \frac{1}{s}\left(\sum_{i=1}^{D} X_{i:n} + \sum_{i=D+1}^{s-1} X_{i:n} + (n-(s-1))X_{s:n}\right). \tag{3.3.3}$$

Furthermore,

$$\hat{\theta}_{121} = \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right) \quad \Rightarrow \quad \sum_{i=1}^{D} X_{i:n} = D\hat{\theta}_{121} - (n-D)T. \tag{3.3.4}$$

Therefore, combining Equations (3.3.3) and (3.3.4) gives us:

$$\hat{\theta}_{122} = \frac{1}{s}\left(D\hat{\theta}_{121} - (n-D)T + \sum_{i=D+1}^{s-1} X_{i:n} + (n-(s-1))X_{s:n}\right). \tag{3.3.5}$$

Conditioning on $D = d$, Theorem (2.1.2) tells us that $X_{d+1:n} \leq X_{d+2:n} \leq \cdots \leq X_{s:n}$ are the order statistics from a complete sample of size $(s-d-1)$ from a standard exponential distribution left-truncated at T. Let $Y_{1:n-d} \leq Y_{2:n-d} \leq \cdots \leq Y_{n-d:n-d}$ represent order statistics that have a standard exponential distribution from a sample of size $(n-d)$. Using Lemma (2.1.3), we will rewrite the expression found in Equation (3.3.5),

$$\begin{aligned}
&\sum_{i=d+1}^{s-1} X_{i:n} + [n-(s-1)]X_{s:n} \\
&= \sum_{i=1}^{(s-1)-d} (Y_{i:n-d} + T) + [n-(s-1)][Y_{s-d:n-d} + T] \\
&= \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [(s-1)-d]T + [n-(s-1)][Y_{s-d:n-d} + T] \\
&= \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d} + [(s-1)-d+n-(s-1)]T \\
&= \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d} + (n-d)T. \qquad (3.3.6)
\end{aligned}$$

Equations (3.3.5) and (3.3.6) allow us to re-write $\hat{\theta}_{122}$:

$$\begin{aligned}
\hat{\theta}_{122} &= \frac{1}{s}\left(d\hat{\theta}_{121} - (n-d)T + \sum_{i=d+1}^{s-1} X_{i:n} + [n-(s-1)]X_{s:n}\right) \\
&= \frac{1}{s}\left(d\hat{\theta}_{121} - (n-d)T + \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d} + (n-d)T\right) \\
&= \frac{1}{s}\left(d\hat{\theta}_{121} + \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d}\right). \qquad (3.3.7)
\end{aligned}$$

As a side note, if $D = s-1$ then the expression $\sum_{i=1}^{(s-1)-d} Y_{i:n-d} = 0$, but the following steps are applicable for deriving the conditional PC probability. Now, we shall explicitly write the expressions for π_{121} and π_{122}:

$$\begin{aligned}
\pi_{121} &= \mathbb{P}_D(2 - \hat{\theta}_{121} \le \hat{\theta}_{122} \le \hat{\theta}_{121}) \\
&= \mathbb{P}_D\left\{2 - \hat{\theta}_{121} \le \frac{1}{s}\left(d\hat{\theta}_{121} + \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d}\right) \le \hat{\theta}_{121}\right\} \\
&= \mathbb{P}_D\left\{2s - \hat{\theta}_{121}(s+d) \le \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d} \le \hat{\theta}_{121}(s-d)\right\} \\
&= \mathbb{P}_D\Bigg\{2s - \frac{s+d}{d}\left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right] \le \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d} \\
&\le \frac{s-d}{d}\left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right]\Bigg\}, \qquad (3.3.8)
\end{aligned}$$

$$\begin{aligned}
\pi_{122} &= \mathbb{P}_D(\hat{\theta}_{121} \le \hat{\theta}_{122} \le 2 - \hat{\theta}_{121}) \\
&= \mathbb{P}_D\left\{\hat{\theta}_{121} \le \frac{1}{s}\left(d\hat{\theta}_{121} + \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d}\right) \le 2 - \hat{\theta}_{121}\right\} \\
&= \mathbb{P}_D\left\{\hat{\theta}_{121}(s-d) \le \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d} \le 2s - \hat{\theta}_{121}(s+d)\right\} \\
&= \mathbb{P}_D\Bigg\{\frac{s-d}{d}\left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right] \le \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d} \\
&\le 2s - \frac{s+d}{d}\left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right]\Bigg\}. \qquad (3.3.9)
\end{aligned}$$

Again, we approach this problem by finding the distribution of the expression in the middle of the inequality shown in Equations (3.3.8) and (3.3.9). Consider:

$$B_{12} := \sum_{i=1}^{(s-1)-d} Y_{i:n-d} + [n-(s-1)]Y_{s-d:n-d}. \tag{3.3.10}$$

To find the distribution of B_{12}, we use the same method as Equation (3.2.8), where if we claim $Y_{0:n-d} = 0$ then we notice that B_{12} is the sum of spacings. That is,

$$B_{12} = \sum_{i=1}^{s-d}((n-d)-i+1)(Y_{i:n-d} - Y_{i-1:n-d}). \tag{3.3.11}$$

We again use Theorem (2.1.1) to claim that $Z_1, Z_2, \ldots, Z_n$ are independent and have standard exponential distributions. In addition, combining Theorem (2.1.1) and Lemma (2.1.1) allows us to claim:

$$B_{12} = \sum_{i=1}^{s-d}((n-d)-i+1)(Y_{i:n-d} - Y_{i-1:n-d}) = \sum_{i=1}^{s-d} Z_i \sim Gamma(s-d,1). \tag{3.3.12}$$

Now, we would like to consider the bounds on B_{12}, which depend on $A_{12} := \sum_{i=1}^{d} X_{i:n}$, which represents the sum of a complete sample of size d that are exponentially distributed but right-truncated at T. Again, we use Proposition (3.2.1) where we set $m = d$ and $L = T$ to obtain:

$$f_{A_{12}|D=d}(a) = \frac{1}{(1-e^{-T})^d}\sum_{i=0}^{d}(-1)^i\binom{d}{i}e^{-iT}\frac{1}{\Gamma(d)}(a-iT)^{d-1}e^{-(a-iT)}I(a > iT). \tag{3.3.13}$$

Furthermore, we have $0 \le A_{12} \le dT$.

The inequalities defining π_{121} and π_{122} as defined in Equations (3.3.8) and (3.3.9) impose additional constraints for A_{12}. For clarity, we denote these bounds explicitly as:

$$L_{12}(A_{12}) := 2s-\frac{1}{d}(A_{12}+(n-d)T)[s+d], \quad U_{12}(A_{12}) := \frac{1}{d}(A_{12}+(n-d)T)[s-d]. \tag{3.3.14}$$

For π_{121}, we need to pay attention to the case when $L_{12}(A_{12}) \leq U_{12}(A_{12})$:

$$\begin{aligned} L_{12}(A_{12}) \leq U_{12}(A_{12}) &\Leftrightarrow 2s - \frac{1}{d}(A_{12} + (n-d)T)[s+d] \leq \frac{1}{d}(A_{12} + (n-d)T)[s-d] \\ &\Leftrightarrow 2s \leq \frac{1}{d}(A_{12} + (n-d)T)[2s] \\ &\Leftrightarrow 1 \leq d(A_{12} + (n-d)T) \\ &\Leftrightarrow d - (n-d)T \leq A_{12}. \end{aligned} \tag{3.3.15}$$

Similarly, for π_{122} we focus on $L_{12}(A_{12}) \geq U_{12}(A_{12})$ which implies $A_{12} \leq d-(n-d)T$. Thus, utilising the total law of probability, π_{121} and π_{122} can be calculated using:

$$\begin{aligned} \pi_{121} &= \mathbb{P}_D\left\{2s - \frac{s+d}{d}\left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right] \leq B_{12} \right. \\ &\left. \leq \frac{s-d}{d}\left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right]\right\} \\ &= \int_{d-(n-d)T}^{dT} \mathbb{P}_D\left(L_{12}(a) \leq B_{12} \leq U_{12}(a) \,|\, A_{12} = a\right) f_{A_{12}|D=d}(a)\, da \\ &= \int_{d-(n-d)T}^{dT} [F_{B_{12}}(U_{12}(a)) - F_{B_{12}}(L_{12}(a))]\, f_{A_{12}|D=d}(a)\, da, \end{aligned} \tag{3.3.16}$$

and

$$\pi_{122} = \mathbb{P}_D\left\{\frac{s-d}{d}\left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right] \leq B_{12}\right.$$

$$\leq 2s - \frac{s+d}{d}\Big[\sum_{i=1}^{d} X_{i:n} + (n-d)T\Big]\Big\}$$
$$= \int_0^{d-(n-d)T} \mathbb{P}_D\left(U_{12}(a) \leq B_{12} \leq L_{12}(a)\right) f_{A_{12}|D=d}(a)\, da$$
$$= \int_0^{d-(n-d)T} [F_{B_{12}}(L_{12}(a)) - F_{B_{12}}(U_{12}(a))]\, f_{A_{12}|D=d}(a)\, da. \tag{3.3.17}$$

Hence, the PC probability associated with the second case is calculated by:

$$\sum_{d=r}^{s-1} \mathbb{P}(D=d)\mathbb{P}_D(|\hat{\theta}_{122} - 1| \leq |\hat{\theta}_{121} - 1|) = \sum_{d=r}^{s-1} \mathbb{P}(D=d)[\pi_{121} + \pi_{122}]. \tag{3.3.18}$$

3.4 Case 3

This is the case where $D \in \{s, \dots, n\}$. Here, the estimators are the same:

$$\hat{\theta}_{131} = \hat{\theta}_{132} = \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right). \tag{3.4.1}$$

Therefore, we can easily say that

$$\sum_{d=s}^{n} \mathbb{P}(D=d)\mathbb{P}_D(|\hat{\theta}_{132} - 1| \leq |\hat{\theta}_{131} - 1|) = \sum_{d=s}^{n} \mathbb{P}(D=d). \tag{3.4.2}$$

3.5 Pitman Closeness Criterion for Comparison of $\hat{\theta}_1$ and $\hat{\theta}_2$

Now that we have derived the exact expressions for the three separate cases, we can combine them to find the PC probability between estimators $\hat{\theta}_1$ and $\hat{\theta}_2$. We have that:

$$\begin{aligned}
\mathbb{P}(|\hat{\theta}_2 - 1| \leq |\hat{\theta}_1 - 1|) &= \sum_{d=0}^{r-1} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{112} - 1| \leq |\hat{\theta}_{111} - 1|) \\
&+ \sum_{d=r}^{s-1} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{122} - 1| \leq |\hat{\theta}_{121} - 1|) \\
&+ \sum_{d=s}^{n} \mathbb{P}(D = d)\mathbb{P}_D(|\hat{\theta}_{132} - 1| \leq |\hat{\theta}_{131} - 1|) \\
&= \sum_{d=0}^{r-1} \mathbb{P}(D = d)[\pi_{111} + \pi_{112}] \\
&+ \sum_{d=r}^{s-1} \mathbb{P}(D = d)[\pi_{121} + \pi_{122}] \\
&+ \sum_{d=s}^{n} \mathbb{P}(D = d),
\end{aligned} \tag{3.5.1}$$

where $\mathbb{P}(D = d)$ is given in Equation (3.1.7). Computational results can be found in Chapter 5.1, and the associated R code can be found in Appendix A.2.

Chapter 4

Comparison of $\hat{\theta}_1$ and $\hat{\theta}_3$

Similar to the previous chapter, we now focus on deriving the Pitman closeness (PC) of two estimators based on Type-II HCS experiments to examine which estimator is Pitman closer to the scale parameter θ of the exponential distribution once you increase the time allocated for the study.

4.1 Estimators and Case Breakdown

Assume a similar setting as the previous chapter, except now we introduce a longer termination time $T^* > T$. Let D^* be a random variable which represents the random number of failures between $(0, T^*]$. Employing an argument parallel to that used for deriving the distribution of D, we find that $D^* = \sum_{i=1}^{n} \mathbf{1}_{X_i}(0, T^*]$ for $X_i \sim \textit{Exponential}(1)$ for $i \in \{1, 2, \ldots, n\}$. This further informs us that $D^* \sim \textit{Binomial}(n, 1 - e^{-T^*})$. Figure 4.1 demonstrates the relationship between D and D^*, assuming that D and D^* are unequal. It is possible for $D = D^*$; this occurs when there are no additional failures observed between T and T^*.

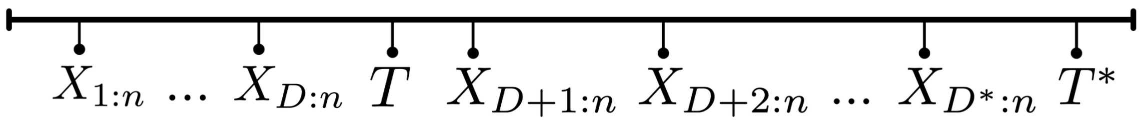

Figure 4.1: Timeline showcasing D and D^* and their relation to times T and T^*.

Let $\hat{\theta}_1$ be the same as Equation (3.1.1):

$$\hat{\theta}_1 = \begin{cases} \frac{1}{r}\left\{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right\} & D = 0, 1, \ldots, r-1 \\ \frac{1}{D}\left\{\sum_{i=1}^{D} X_{i:n} + (n-D)T\right\} & D = r, r+1, \ldots, n, \end{cases} \tag{4.1.1}$$

and $\hat{\theta}_3$ be the respective MLE for θ under the termination times $\max\{X_{r:n}, T^*\}$. Again, the result from Childs et al. [11] suggests this third estimator is of the form:

$$\hat{\theta}_3 = \begin{cases} \frac{1}{r}\left\{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right\} & D^* = 0, 1, \ldots, r-1, \\ \frac{1}{D^*}\left\{\sum_{i=1}^{D^*} X_{i:n} + (n-D^*)T^*\right\} & D^* = r, r+1, \ldots, n. \end{cases} \tag{4.1.2}$$

The expressions for $\hat{\theta}_1$ and $\hat{\theta}_3$ depend on the values of D and D^*, hence we consider the following cases:

Case 1: $D, D^* \in \{0, 1, \ldots, r-1\}$,

Case 2: $D, D^* \in \{r, r+1, \ldots, n\}$,

Case 3: $D \in \{0, 1, \ldots, r-1\}$ and $D^* \in \{r, r+1, \ldots, n\}$.

In case 1, we have that:

$$\hat{\theta}_{211} := \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right), \quad \hat{\theta}_{213} := \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right). \tag{4.1.3}$$

Furthermore, for case 2, we have that:

$$\hat{\theta}_{221} := \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right), \quad \hat{\theta}_{223} := \frac{1}{D^*}\left(\sum_{i=1}^{D^*} X_{i:n} + (n-D^*)T^*\right). \tag{4.1.4}$$

Finally, for case 3, we have that:

$$\hat{\theta}_{231} := \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right), \quad \hat{\theta}_{233} := \frac{1}{D^*}\left(\sum_{i=1}^{D^*} X_{i:n} + (n-D^*)T^*\right). \tag{4.1.5}$$

For convenient notation, allow $\mathbb{P}_{D,D^*}(\cdot) := \mathbb{P}\{\cdot | D = d, D^* = d^*\}$. We also borrow the earlier notation where $\mathbb{P}_D(\cdot) := \mathbb{P}\{\cdot | D = d\}$. Then, utilising the total law of probability, we can compute the PC probability via the following formula:

$$\begin{aligned}\mathbb{P}(|\hat{\theta}_3 - 1| \leq |\hat{\theta}_1 - 1|) = &\sum_{d=0}^{r-1} \mathbb{P}(D = d) \sum_{d^*=d}^{r-1} \mathbb{P}_D(D^* = d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{213} - 1| \leq |\hat{\theta}_{211} - 1|) \\ &+ \sum_{d=r}^{n} \mathbb{P}(D = d) \sum_{d^*=d}^{n} \mathbb{P}_D(D^* = d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{223} - 1| \leq |\hat{\theta}_{221} - 1|) \\ &+ \sum_{d=0}^{r-1} \mathbb{P}(D = d) \sum_{d^*=r}^{n} \mathbb{P}_D(D^* = d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{233} - 1| \leq |\hat{\theta}_{231} - 1|).\end{aligned} \tag{4.1.6}$$

Again, $\mathbb{P}(D = d)$ can be found through Equation (3.1.7). It is worth recalling that:

$$\mathbb{P}(D = d) = \binom{n}{d}(1 - e^{-T})^d (e^{-T})^{n-d}, \tag{4.1.7}$$

but calculating $\mathbb{P}_D(D^* = d^*)$ is more challenging. Conditioning on $D = d$, let D' represent the number of lifetimes between $X_{d:n}$ and $X_{D^*:n}$. That is, $D' = D^* - d$. With

$(n-d)$ lifetimes remaining, the probability of a "success" (i.e., the probability that a lifetime from a standard exponential distribution lies between T and T^*, conditional on exceeding T) is:

$$\begin{aligned}\mathbb{P}(T < X_i \le T^* \mid X_i > T) &= \frac{\mathbb{P}(T < X_i \le T^*)}{\mathbb{P}(X_i > T)} \\ &= \frac{\mathbb{P}(X_i \le T^*) - \mathbb{P}(X_i < T)}{\mathbb{P}(X_i > T)} \\ &= \frac{(1-e^{-T^*}) - (1-e^{-T^*})}{e^{-T}} \\ &= 1 - e^{-(T^*-T)}. \end{aligned} \tag{4.1.8}$$

Thus, we claim that $D' \mid D = d \sim Binomial(n-d, 1-e^{-(T^*-T)})$, and we make the claim that $\mathbb{P}_D(D^* = d^*) = \mathbb{P}_D(D' = d^* - d)$. Therefore,

$$\mathbb{P}_D(D' = d') = \binom{n-d}{d'}(1-e^{-(T^*-T)})^{d'}(e^{-(T^*-T)})^{(n-d)-d'}. \tag{4.1.9}$$

Similar to before, we will compute the following cases for the next subchapters:

Case 1: $\sum_{d=0}^{r-1} \mathbb{P}(D = d) \sum_{d^*=d}^{r-1} \mathbb{P}_D(D^* = d^*) \mathbb{P}_{D,D^*}(|\hat{\theta}_{213} - 1| \le |\hat{\theta}_{211} - 1|)$,

Case 2: $\sum_{d=r}^{n} \mathbb{P}(D = d) \sum_{d^*=d}^{n} \mathbb{P}_D(D^* = d^*) \mathbb{P}_{D,D^*}(|\hat{\theta}_{223} - 1| \le |\hat{\theta}_{221} - 1|)$,

Case 3: $\sum_{d=0}^{r-1} \mathbb{P}(D = d) \sum_{d^*=r}^{n} \mathbb{P}_D(D^* = d^*) \mathbb{P}_{D,D^*}(|\hat{\theta}_{233} - 1| \le |\hat{\theta}_{231} - 1|)$.

4.2 Case 1

If we have that $D, D^* \in \{0, 1, \ldots, r-1\}$, then it is the case that:

$$\hat{\theta}_{211} = \hat{\theta}_{213} = \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right). \tag{4.2.1}$$

These are the same estimators, hence:

$$\begin{aligned}
&\sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=d}^{r-1} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{213}-1| \leq |\hat{\theta}_{211}-1|) \\
&= \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=d}^{r-1} \mathbb{P}_D(D^*=d^*) \\
&= \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d'=1}^{r-d-1} \mathbb{P}_D(D'=d').
\end{aligned} \tag{4.2.2}$$

4.3 Case 2

If we have that $D, D^* \in \{r, r+1, \ldots, n\}$ then:

$$\hat{\theta}_{221} = \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right), \quad \hat{\theta}_{223} = \frac{1}{D^*}\left(\sum_{i=1}^{D^*} X_{i:n} + (n-D^*)T^*\right). \tag{4.3.1}$$

We need to consider two sub-cases: either $D = D^*$ or $D < D^*$.

4.3.1 Case 2.1

If $D = D^*$ then we have the following:

$$\sum_{i=1}^{D} X_{i:n} = D\hat{\theta}_{221} - (n-D)T. \tag{4.3.2}$$

Therefore,

$$\begin{aligned}
\hat{\theta}_{223} &= \frac{1}{D^*}\left(\sum_{i=1}^{D^*} X_{i:n} + (n-D^*)T^*\right) \\
&= \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T^*\right) \\
&= \frac{1}{D}\left(D\hat{\theta}_{221} - (n-D)T + (n-D)T^*\right) \\
&= \hat{\theta}_{221} + \frac{(n-D)(T^*-T)}{D}. \qquad (4.3.3)
\end{aligned}$$

We then have the following computation:

$$\begin{aligned}
\mathbb{P}_{D,D^*}(|\hat{\theta}_{223}-1| \leq |\hat{\theta}_{221}-1|) &= \mathbb{P}_{D,D^*}(2-\hat{\theta}_{221} \leq \hat{\theta}_{223} \leq \hat{\theta}_{221}) \\
&\quad + \mathbb{P}_{D,D^*}(\hat{\theta}_{221} \leq \hat{\theta}_{223} \leq 2-\hat{\theta}_{221}) \\
&= \pi_{2211} + \pi_{2212}, \quad \text{say,} \qquad (4.3.4)
\end{aligned}$$

where π_{2211} and π_{2212} are used for short-hand. For further convenience, let $M := \frac{(n-D)(T^*-T)}{D}$. We will try to solve for π_{2211} and π_{2212} separately.

$$\begin{aligned}
\pi_{2211} &= \mathbb{P}_{D,D^*}\left\{2-\hat{\theta}_{221} \leq \hat{\theta}_{221} + M \leq \hat{\theta}_{221}\right\} \\
&= \mathbb{P}_{D,D^*}\left\{2-2\hat{\theta}_{221} \leq M \leq 0\right\}. \qquad (4.3.5)
\end{aligned}$$

Since $n > D$ and $T^* > T$, it is impossible for $M \leq 0$ and therefore $\pi_{2211} = 0$. For the second probability in Equation (4.3.4):

$$\begin{aligned}
\pi_{2212} &= \mathbb{P}_{D,D^*}\left\{\hat{\theta}_{221} \leq \hat{\theta}_{221} + M \leq 2-\hat{\theta}_{221}\right\} \\
&= \mathbb{P}_{D,D^*}\left\{0 \leq M \leq 2-2\hat{\theta}_{221}\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}_{D,D^*}\left\{\frac{M-2}{-2} \geq \hat{\theta}_{221}\right\} \\
&= \mathbb{P}_{D,D^*}\left\{1 - \frac{M}{2} \geq \frac{\sum_{i=1}^{d} X_{i:n} + (n-d)T}{d}\right\} \\
&= \mathbb{P}_{D,D^*}\left\{d\left(1 - \frac{M}{2}\right) - (n-d)T \geq \sum_{i=1}^{d} X_{i:n}\right\}. \qquad (4.3.6)
\end{aligned}
$$

From Theorem (2.1.2), we have that $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{d:n}$ represents a complete sample of size d of i.i.d. exponential random variables that are right-truncated at T. Note that we derived the pdf of $A_{12} := \sum_{i=1}^{d} X_{i:n}$ earlier in first comparison in Chapter 3 as Equation (3.3.13). Thus, let:

$$
U_{221}(A_{12}, d, d^*) = d\left(1 - \frac{M}{2}\right) - (n-d)T, \quad M = \frac{(n-D)(T^* - T)}{D}. \qquad (4.3.7)
$$

Note that A_{12} does not depend on $D^* = d^*$, hence that extra condition is omitted when writing the pdf. Therefore, the probability in Equation (4.3.6) can be calculated using:

$$
\pi_{2212} = \int_0^{U_{221}(A_{12},d,d^*)} f_{A_{12}|D=d}(a)da. \qquad (4.3.8)
$$

4.3.2 Case 2.2

Here, we aim to compute:

$$
\begin{aligned}
\mathbb{P}_{D,D^*}(|\hat{\theta}_{223} - 1| \leq |\hat{\theta}_{221} - 1|) &= \mathbb{P}_{D,D^*}\left\{2 - \hat{\theta}_{221} \leq \hat{\theta}_{223} \leq \hat{\theta}_{221}\right\} \\
&\quad + \mathbb{P}_{D,D^*}\left\{\hat{\theta}_{221} \leq \hat{\theta}_{223} \leq 2 - \hat{\theta}_{221}\right\} \\
&= \pi_{2221} + \pi_{2222}, \quad \text{say}, \qquad (4.3.9)
\end{aligned}
$$

where π_{2221} and π_{2222} are denoted for convenience. If we condition on $D = d, D^* = d^*$ then for $d < d^*$ then we have the following:

$$\sum_{i=1}^{d*} X_{i:n} = \sum_{i=1}^{d} X_{i:n} + \sum_{i=d+1}^{d^*} X_{i:n}, \tag{4.3.10}$$

thus,

$$\sum_{i=1}^{d} X_{i:n} = d\hat{\theta}_{221} - (n-d)T. \tag{4.3.11}$$

This implies that:

$$\begin{aligned} \hat{\theta}_{223} &= \frac{1}{d^*}\left(\sum_{i=1}^{d^*} X_{i:n} + (n-d^*)T^*\right) \\ &= \frac{1}{d^*}\left(d\hat{\theta}_{221} - (n-d)T + \sum_{i=d+1}^{d^*} X_{i:n} + (n-d^*)T^*\right). \end{aligned} \tag{4.3.12}$$

Notice $X_{d+1:n} \leq X_{d+2:n} \leq \cdots \leq X_{d^*:n}$ are the order statistics from a sample of size (d^*-d) from the exponential distribution doubly truncated between $(T, T^*]$. Suppose:

$$\begin{aligned} \sum_{i=d+1}^{d^*} X_{i:n} &= \sum_{i=d+1}^{d^*} (X_{i:n} - T) + (d^* - d)T \\ &= \sum_{i=1}^{d^*-d} Y_{i:d^*-d} + (d^* - d)T, \end{aligned} \tag{4.3.13}$$

where $Y_{1:d^*-d} \leq Y_{2:d^*-d} \leq \cdots \leq Y_{d^*-d:d^*-d}$ represent a complete sample of order statistics of the exponential distribution right-truncated at $T' = T^* - T$ of size $(d^* - d)$. We now have:

$$\hat{\theta}_{223} = \frac{1}{d^*}\left(d\hat{\theta}_{221} - (n-d)T + \left[\sum_{i=1}^{d^*-d} Y_{i:d^*-d} + (d^* - d)T\right] + (n-d^*)T^*\right)$$

$$= \frac{1}{d^*}\left(d\hat{\theta}_{221} + \sum_{i=1}^{d^*-d} Y_{i:d^*-d} + (n-d^*)(T^*-T)\right). \tag{4.3.14}$$

Since we have $\hat{\theta}_{221} = \frac{1}{d}\left(\sum_{i=1}^{d} X_{i:n} + (n-d)T\right)$, we will want to condition on $A_{12} := \sum_{i=1}^{d} X_{i:n}$. As mentioned in the previous case, the pdf for A_{12} was found in Equation (3.3.13) where its bounds are $0 \leq A_{12} \leq dT$, so we have:

$$\begin{aligned}
\pi_{2221} &= \mathbb{P}_{D,D^*}\left\{2 - \hat{\theta}_{221} \leq \frac{1}{d^*}\left(d\hat{\theta}_{221} + \sum_{i=1}^{d^*-d} Y_{i:d^*-d} + (n-d^*)(T^*-T)\right) \leq \hat{\theta}_{221}\right\} \\
&= \mathbb{P}_{D,D^*}\Big\{2d^* - \hat{\theta}_{221}d^* - d\hat{\theta}_{221} - (n-d^*)(T^*-T) \\
&\leq \sum_{i=1}^{d^*-d} Y_{i:d^*-d} \leq \hat{\theta}_{221}d^* - d\hat{\theta}_{221} - (n-d^*)(T^*-T)\Big\} \\
&= \mathbb{P}_{D,D^*}\Big\{2d^* - \hat{\theta}_{221}(d^*+d) - (n-d^*)(T^*-T) \\
&\leq \sum_{i=1}^{d^*-d} Y_{i:d^*-d} \leq \hat{\theta}_{221}(d^*-d) - (n-d^*)(T^*-T)\Big\} \\
&= \mathbb{P}_{D,D^*}\Big\{2d^* - \frac{(A_{12}+(n-d)T)}{d}(d^*+d) - (n-d^*)(T^*-T) \\
&\leq \sum_{i=1}^{d^*-d} Y_{i:d^*-d} \leq \frac{(A_{12}+(n-d)T)}{d}(d^*-d) - (n-d^*)(T^*-T)\Big\}.
\end{aligned} \tag{4.3.15}$$

Furthermore,

$$\begin{aligned}
\pi_{2222} &= \mathbb{P}_{D,D^*}\left\{\hat{\theta}_{221} \leq \frac{1}{d^*}\left(d\hat{\theta}_{221} + \sum_{i=1}^{d^*-d} Y_{i:d^*-d} + (n-d^*)(T^*-T)\right) \leq 2 - \hat{\theta}_{221}\right\} \\
&= \mathbb{P}_{D,D^*}\Big\{\frac{(A_{12}+(n-d)T)}{d}(d^*-d) - (n-d^*)(T^*-T) \\
&\leq \sum_{i=1}^{d^*-d} Y_{i:d^*-d} \leq 2d^* - \frac{(A_{12}+(n-d)T)}{d}(d^*+d) - (n-d^*)(T^*-T)\Big\}.
\end{aligned} \tag{4.3.16}$$

We are now interested in finding the pdf for the middle expression in the inequalities shown in Equations (4.3.15) and (4.3.16), which we will denote as $B_{22} := \sum_{i=1}^{d^*-d} Y_{i:d^*-d}$. Conditioning on $D = d, D^* = d^*$, this density function can also be found using Proposition (3.2.1) where $m = d^* - d$ and $L = T' = T^* - T$. Thus, we have:

$$f_{B_{22}|D=d,D^*=d^*}(b) = \frac{1}{(1-e^{-T'})^{(d^*-d)}} \sum_{i=0}^{d^*-d} \binom{d^*-d}{i} (-1)^i e^{-T'i} \times \frac{1}{\Gamma(d^*-d)} (b - T'i)^{(d^*-d)-1} e^{-(b-T'i)} I(b > T'i). \tag{4.3.17}$$

Now, we will define the bounds on B_{22} found in Equations (4.3.15) and (4.3.16):

$$L_{222}(A_{12}, d, d^*) := 2d^* - \frac{(A_{12} + (n-d)T)}{d}(d^* + d) - (n - d^*)(T^* - T), \tag{4.3.18}$$

$$U_{222}(A_{12}, d, d^*) := \frac{(A_{12} + (n-d)T)}{d}(d^* - d) - (n - d^*)(T^* - T). \tag{4.3.19}$$

Note that,

$$\begin{aligned} L_{222}(A_{12}, d, d^*) \leq U_{222}(A_{12}, d, d^*) &\Leftrightarrow 2d^* - \frac{(A_{12} + (n-d)T)}{d}(d^* + d) \\ &\leq \frac{(A_{12} + (n-d)T)}{d}(d^* - d) \\ &\Leftrightarrow 2d^* \leq \frac{(A_{12} + (n-d)T)}{d}(2d^*) \\ &\Leftrightarrow d \leq (A_{12} + (n-d)T) \\ &\Leftrightarrow d - (n-d)T \leq A_{12}. \end{aligned} \tag{4.3.20}$$

Similarly, $U_{222}(A_{12}, d, d^*) \leq L_{222}(A_{12}, d, d^*)$ tells us that $A_{12} \leq d-(n-d)T$. Therefore, utilising the total law of probability, we claim that:

$$\begin{aligned}
\pi_{2221} &= \int_{d-(n-d)T}^{dT} \mathbb{P}(L_{222}(a, d, d^*) \leq B_{22} \leq U_{222}(a, d, d^*) \,|\, A_{12} = a, D = d, D^* = d) \\
&\times f_{A_{12}|D=d}(a)\, da \\
&= \int_{d-(n-d)T}^{dT} [F_{B_{22}|(A_{12}=a,D=d,D^*=d^*)}(U_{222}(a, d, d^*)) - F_{B_{22}|(A_{12}=a,D=d,D^*=d^*)}(L_{222}(a, d, d^*))] \\
&\times f_{A_{12}|D=d}(a)\, da,
\end{aligned} \tag{4.3.21}$$

and similarly,

$$\begin{aligned}
\pi_{2222} &= \int_{0}^{d-(n-d)T} \mathbb{P}(U_{222}(a, d, d^*) \leq B_{22} \leq L_{222}(a, d, d^*) \,|\, A_{12} = a, D = d, D^* = d^*) \\
&\times f_{A_{12}|D=d}(a)\, da \\
&= \int_{0}^{d-(n-d)T} [F_{B_{22}|(A_{12}=a,D=d,D^*=d)}(L_{222}(a, d, d^*)) - F_{B_{22}|(A_{12}=a,D=d,D^*=d)}(U_{222}(a, d, d^*))] \\
&\times f_{A_{12}|D=d}(a)\, da.
\end{aligned} \tag{4.3.22}$$

Therefore, the PC probability corresponding to the second case is:

$$\begin{aligned}
&\sum_{d=r}^{n} \mathbb{P}(D = d) \sum_{d^*=d}^{n} \mathbb{P}_D(D^* = d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{223} - 1| \leq |\hat{\theta}_{221} - 1|) \\
&= \sum_{d=r}^{n} \mathbb{P}(D = d) \left(\mathbb{P}_D(D^* = d)\pi_{221} + \sum_{d^*=d+1}^{n} \mathbb{P}_D(D^* = d^*)[\pi_{2221} + \pi_{2222}] \right) \\
&= \sum_{d=r}^{n} \mathbb{P}(D = d) \left(\mathbb{P}_D(D' = 0)\pi_{221} + \sum_{d'=1}^{n-d} \mathbb{P}_D(D' = d')[\pi_{2221} + \pi_{2222}] \right).
\end{aligned} \tag{4.3.23}$$

4.4 Case 3

In this case, $D \in \{0, 1, \ldots, r-1\}$ and $D^* \in \{r, r+1, \ldots, n\}$. Hence, we will compare the following estimators:

$$\hat{\theta}_{231} = \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right), \;\; \hat{\theta}_{233} = \frac{1}{D^*}\left(\sum_{i=1}^{D^*} X_{i:n} + (n-D^*)T^*\right). \tag{4.4.1}$$

Again, we consider two sub-cases, either $D^* = r$ or $D^* > r$.

4.4.1 Case 3.1

Suppose that $D^* = r$. Since D^* is the number of failures observed before time T^*, we have $T^* \geq X_{r:n}$. Therefore, we have that:

$$\begin{aligned}
T^* \geq X_{r:n} &\Leftrightarrow (n-r)T^* \geq (n-r)X_{r:n} \\
&\Leftrightarrow \sum_{i=1}^{r} X_{i:n} + (n-r)T^* \geq \sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n} \\
&\Leftrightarrow \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)T^*\right) \geq \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right) \\
&\Leftrightarrow \hat{\theta}_{233} \geq \hat{\theta}_{231}. \qquad (4.4.2)
\end{aligned}$$

Deriving the conditional PC probability, we obtain:

$$\begin{aligned}
\pi_{231} &:= \mathbb{P}_{D,D^*}(|\hat{\theta}_{233} - 1| \leq |\hat{\theta}_{231} - 1|) \\
&= \mathbb{P}_{D,D^*}\{2 - \hat{\theta}_{231} \leq \hat{\theta}_{233} \leq \hat{\theta}_{231}\} + \mathbb{P}_{D,D^*}\{\hat{\theta}_{231} \leq \hat{\theta}_{233} \leq 2 - \hat{\theta}_{231}\} \\
&= 0 + \mathbb{P}_{D,D^*}(\hat{\theta}_{233} \leq 2 - \hat{\theta}_{231}), \qquad (4.4.3)
\end{aligned}$$

where the last line arises from the fact that $\hat{\theta}_{233} \geq \hat{\theta}_{231}$. For convenience, let us denote $\pi_{231} := \mathbb{P}_{D,D^*}(\hat{\theta}_{233} \leq 2 - \hat{\theta}_{231})$. Now, note that:

$$
\begin{aligned}
\hat{\theta}_{233} \leq 2 - \hat{\theta}_{231} &\Leftrightarrow \frac{1}{r}\left(\sum_{i=1}^{D^*} X_{i:n} + (n - D^*)T^*\right) \leq 2 - \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right) \\
&\Leftrightarrow \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)T^*\right) \leq 2 - \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right) \\
&\Leftrightarrow \frac{2}{r}\sum_{i=1}^{r} X_{i:n} + \left(\frac{n-r}{r}\right)X_{r:n} \leq 2 - \left(\frac{n-r}{r}\right)T^* \\
&\Leftrightarrow \sum_{i=1}^{r} X_{i:n} + \left(\frac{n-r}{2}\right)X_{r:n} \leq r - \left(\frac{n-r}{2}\right)T^* \\
&\Leftrightarrow \sum_{i=1}^{r-1} X_{i:n} + \left(1 + \frac{n-r}{2}\right)X_{r:n} \leq r - \left(\frac{n-r}{2}\right)T^* \\
&\Leftrightarrow \sum_{i=1}^{d} X_{i:n} + \sum_{i=d+1}^{r-1} X_{i:n} + \left(1 + \frac{n-r}{2}\right)X_{r:n} \leq r - \left(\frac{n-r}{2}\right)T^*. \qquad (4.4.4)
\end{aligned}
$$

When deriving the conditional PC probability for this case, we must consider two things. First, by Theorem (2.1.2) we have that $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{d:n}$ represent the order statistics from a complete sample of size d drawn from an exponential distribution that is right-truncated at T. Furthermore, the pdf of $A_{12} := \sum_{i=1}^{d} X_{i:n}$ has been calculated before in Equation (3.3.13) and its bounds are $0 \leq A_{12} \leq dT$.

Second, $X_{d+1:n} \leq X_{d+2:n} \leq \cdots \leq X_{r:n}$ are the order statistics doubly truncated between T and T^* from a complete sample of size $(r-d)$. Then, Lemma (2.1.4) allows us to claim that $X_{j:n} - T$ represents an order statistic that is only right-truncated at $T'' = T^* - T$ for $j = d+1, d+2, \ldots, r$. Hence, let $Y_{1:r-d} \leq Y_{2:r-d} \leq \cdots \leq Y_{r-d:r-d}$ represent these order statistics that are right-truncated at T'' from a complete sample of size

$(r-d)$, and consider the following:

$$
\begin{aligned}
&\sum_{i=d+1}^{r-1} X_{i:n} + \left(1 + \frac{n-r}{2}\right) X_{r:n} \\
&= \sum_{i=d+1}^{r-1} X_{i:n} + \left(1 + \frac{n-r}{2}\right) X_{r:n} - (r-d-1)T + (r-d-1)T \\
&= \sum_{i=d+1}^{r-1} (X_{i:n} - T) + \left(1 + \frac{n-r}{2}\right) X_{r:n} + (r-d-1)T \\
&= \sum_{i=d+1}^{r-1} (X_{i:n} - T) + \left(1 + \frac{n-r}{2}\right) X_{r:n} + (r-d-1)T - \left(1 + \frac{n-r}{2}\right) T + \left(1 + \frac{n-r}{2}\right) T \\
&= \sum_{i=d+1}^{r-1} (X_{i:n} - T) + \left(1 + \frac{n-r}{2}\right) (X_{r:n} - T) + (r-d-1)T + \left(1 + \frac{n-r}{2}\right) T \\
&= \sum_{i=1}^{r-d-1} Y_{i:r-d} + \left(1 + \frac{n-r}{2}\right) Y_{r-d:r-d} + (r-d-1)T + \left(1 + \frac{n-r}{2}\right) T. \qquad (4.4.5)
\end{aligned}
$$

So instead, we can focus on:

$$
\begin{aligned}
\hat{\theta}_{233} \leq 2 - \hat{\theta}_{231} &\Leftrightarrow \sum_{i=1}^{d} X_{i:n} + \sum_{i=d+1}^{r-1} X_{i:n} + \left(1 + \frac{n-r}{2}\right) X_{r:n} \leq r - \left(\frac{n-r}{2}\right) T^* \\
&\Leftrightarrow A_{12} + \sum_{i=1}^{r-d-1} Y_{i:r-d} + \left(1 + \frac{n-r}{2}\right) Y_{r-d:r-d} + (r-d-1)T \\
&+ \left(1 + \frac{n-r}{2}\right) T \leq r - \left(\frac{n-r}{2}\right) T^* \\
&\Leftrightarrow A_{12} + \sum_{i=1}^{r-d-1} Y_{i:r-d} + \left(1 + \frac{n-r}{2}\right) Y_{r-d:r-d} \\
&\leq r - \left(\frac{n-r}{2}\right) T^* - (r-d-1)T - \left(1 + \frac{n-r}{2}\right) T. \qquad (4.4.6)
\end{aligned}
$$

To proceed further, we would be interested in computing the pdf for $B_{231} := \sum_{i=1}^{r-d-1} Y_{i:r-d} + \omega Y_{r-d:r-d}$ where $\omega \in \mathbb{R}$. We next present a proposition corresponding to Equation (65) in Davies [13]:

Proposition 4.4.1. ***Sum of Right-Truncated Ordered Exponential with Weighted Maximum***. *Let $\tau_i = L(\omega + i)$, $\theta_i = \frac{\omega+i}{i+1}$ and J represents the lower incomplete gamma function:*

$$J(a,x) := \frac{1}{\Gamma(a)} \int_0^x t^{a-1} e^{-t} dt. \tag{4.4.7}$$

Now, suppose $X_{1:n} \leq \cdots \leq X_{n:n}$ be the order statistics obtained from a sample of size n from the Exponential(1) distribution. If the obtained data are truncated on the right at L, and exactly m of the X_i's are less than or equal to L, then, the first m order statistics $X_{1:n} \leq \cdots \leq X_{m:n}$ form a complete random sample of size m from the right-truncated exponential distribution at L. Then, for $\omega \in \mathbb{R}$, the pdf of $Z' := \sum_{i=1}^{m-1} X_{i:n} + \omega X_{m:n}$ is:

$$\begin{aligned} f_{Z^*}(z) = \frac{m}{(1-e^{-L})^m} \sum_{i=0}^{m-1} (-1)^i \binom{m-1}{i} \Bigg[& \frac{1}{\theta_i(i+1)} e^{-z/\theta_i} \frac{1}{\left(1-\frac{1}{\theta_i}\right)^{(m-1)}} J\left(m-1, \left(1-\frac{1}{\theta_i}\right) z\right) \\ & - \frac{e^{-L(i+1)}}{\theta_i(i+1)} e^{-(z-\tau_i)/\theta_i} \frac{1}{\left(1-\frac{1}{\theta_i}\right)^{m-1}} J\left(m-1, \left(1-\frac{1}{\theta_i}\right)(z-\tau_i)\right) \Bigg], \end{aligned} \tag{4.4.8}$$

Furthermore, $0 \leq Z' \leq (m-1+\omega)L$.

Using Proposition (4.4.1), we can substitute $m = r - d$ and $L = T'' = T^* - T$ to obtain the pdf for B_{231}:

$$\begin{aligned} f_{B_{231}}(b) = & \frac{r-d}{(1-e^{-T''})^{(r-d)}} \sum_{i=0}^{r-d-1} (-1)^i \binom{r-d-1}{i} \\ & \times \left[\frac{1}{\theta_i(i+1)} e^{-b/\theta_i} \frac{1}{\left(1-\frac{1}{\theta_i}\right)^{(r-d-1)}} J\left(r-d-1, \left(1-\frac{1}{\theta_i}\right) b\right) \right. \end{aligned}$$

$$- \frac{e^{-T''(i+1)}}{\theta_i(i+1)} e^{-(b-\tau_i)/\theta_i} \frac{1}{\left(1-\frac{1}{\theta_i}\right)^{r-d-1}} J\left(r-d-1, \left(1-\frac{1}{\theta_i}\right)(b-\tau_i)\right)\Bigg]. \tag{4.4.9}$$

Here, $\omega = 1 + \frac{n-r}{2}$ and $\tau_i = T''(\omega + i)$. Returning to deriving the conditional PC probability shown in Equation (4.4.6):

$$\begin{aligned}
\hat{\theta}_{233} \leq 2 - \hat{\theta}_{231} \Leftrightarrow A_{12} + \sum_{i=1}^{r-d-1} Y_{i:r-d} + \left(1 + \frac{n-r}{2}\right) Y_{r-d:r-d} \\
\leq r - \left(\frac{n-r}{2}\right) T^* - (r-d-1)T - \left(1 + \frac{n-r}{2}\right) T \\
\Leftrightarrow B_{231} \leq r - \left(\frac{n-r}{2}\right) T^* - (r-d-1)T - \left(1 + \frac{n-r}{2}\right) T - A_{12}. \quad (4.4.10)
\end{aligned}$$

For notational convenience, let:

$$U_{231}(A_{12}) := r - \left(\frac{n-r}{2}\right) T^* - (r-d-1)T - \left(1 + \frac{n-r}{2}\right) T - A_{12}. \tag{4.4.11}$$

Then we have that,

$$\begin{aligned}
\pi_{231} &= \mathbb{P}_{D,D^*}\left(B_{231} \leq U_{231}(A_{12})\right) \\
&= \int_0^{dT} \mathbb{P}_{D,D^*}(B_{231} \leq U_{231}(a) | A_{12} = a) f_{A_{12}|D=d,D^*=d}(a)\, da \\
&= \int_0^{dT} F_{B_{231}|A_{12}=a}(U_{231}(a)) f_{A_{12}|D=d,D^*=d}(a)\, da \\
&= \int_0^{dT} F_{B_{231}}(U_{231}(a)) f_{A_{12}|D=d}(a)\, da. \qquad (4.4.12)
\end{aligned}$$

4.4.2 Case 3.2

Suppose that $D^* > r$. Then, one way to rewrite the conditional PC probability is:

$$\begin{aligned}\mathbb{P}_{D,D^*}(|\hat{\theta}_{233} - 1| \leq |\hat{\theta}_{231} - 1|) &= \mathbb{P}_{D,D^*}(2 - \hat{\theta}_{231} \leq \hat{\theta}_{233} \leq \hat{\theta}_{231}) \\ &\quad + \mathbb{P}_{D,D^*}(\hat{\theta}_{231} \leq \hat{\theta}_{233} \leq 2 - \hat{\theta}_{231}) \\ &= \pi_{2321} + \pi_{2322}, \ \text{say}, \end{aligned} \tag{4.4.13}$$

where π_{2321} and π_{2322} are introduced for notational convenience. First, we will focus on evaluating π_{2321}:

$$\begin{aligned}\pi_{2321} &= \mathbb{P}_{D,D^*}\Bigg\{2 - \frac{1}{r}\Big[\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Big] \leq \frac{1}{d^*}\Big[\sum_{i=1}^{d^*} X_{i:n} + (n-d^*)T\Big] \\ &\leq \frac{1}{r}\Big[\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Big]\Bigg\} \\ &= \mathbb{P}_{D,D^*}\Bigg\{2 - \frac{1}{r}\Big[\sum_{i=1}^{r-1} X_{i:n} + (n-r+1)X_{r:n}\Big] - \Big(\frac{1}{d^*}\sum_{i=1}^{r} X_{i:n} + \frac{(n-d^*)T}{d^*}\Big) \\ &\leq \frac{1}{d^*}\Big[\sum_{i=r+1}^{d^*} X_{i:n}\Big] \leq \frac{1}{r}\Big[\sum_{i=1}^{r-1} X_{i:n} + (n-r+1)X_{r:n}\Big] - \Big(\frac{1}{d^*}\sum_{i=1}^{r} X_{i:n} + \frac{(n-d^*)T}{d^*}\Big)\Bigg\} \\ &= \mathbb{P}_{D,D^*}\Bigg\{2 - \Big(\frac{1}{r} + \frac{1}{d^*}\Big)\sum_{i=1}^{r-1} X_{i:n} - \Big(\frac{n-r+1}{r} + \frac{1}{d^*}\Big)X_{r:n} - \frac{(n-d^*)T}{d^*} \\ &\leq \frac{1}{d^*}\Big[\sum_{i=r+1}^{d^*} X_{i:n}\Big] \leq \Big(\frac{1}{r} - \frac{1}{d^*}\Big)\sum_{i=1}^{r-1} X_{i:n} + \Big(\frac{n-r+1}{r} - \frac{1}{d^*}\Big)X_{r:n} - \frac{(n-d^*)T}{d^*}\Bigg\}. \end{aligned} \tag{4.4.14}$$

Consider the following:

$$\frac{1}{d^*}\sum_{i=r+1}^{d^*} X_{i:n} = \frac{1}{d^*}\sum_{i=r+1}^{d^*}(X_{i:n} - X_{r:n}) + \frac{d^* - r}{d^*}X_{r:n}. \tag{4.4.15}$$

Thus, we have that:

$$\begin{aligned}
\pi_{2321} &= \mathbb{P}_{D,D^*}\Bigg\{2-\left(\frac{1}{r}+\frac{1}{d^*}\right)\sum_{i=1}^{r-1}X_{i:n}-\left(\frac{n-r+1}{r}+\frac{1}{d^*}\right)X_{r:n}-\frac{(n-d^*)T}{d^*}\\
&\leq \frac{1}{d^*}\sum_{i=r+1}^{d^*}(X_{i:n}-X_{r:n})+\frac{d^*-r}{d^*}X_{r:n}\\
&\leq \left(\frac{1}{r}-\frac{1}{d^*}\right)\sum_{i=1}^{r-1}X_{i:n}+\left(\frac{n-r+1}{r}-\frac{1}{d^*}\right)X_{r:n}-\frac{(n-d^*)T}{d^*}\Bigg\}\\
&= \mathbb{P}_{D,D^*}\Bigg\{2-\left(\frac{1}{r}+\frac{1}{d^*}\right)\sum_{i=1}^{r-1}X_{i:n}-\left(\frac{n-r+1}{r}+\frac{d^*-r+1}{d^*}\right)X_{r:n}-\frac{(n-d^*)T}{d^*}\\
&\leq \frac{1}{d^*}\sum_{i=r+1}^{d^*}(X_{i:n}-X_{r:n})\\
&\leq \left(\frac{1}{r}-\frac{1}{d^*}\right)\sum_{i=1}^{r-1}X_{i:n}+\left(\frac{n-r+1}{r}-\frac{d^*-r+1}{d^*}\right)X_{r:n}-\frac{(n-d^*)T}{d^*}\Bigg\}. \qquad (4.4.16)
\end{aligned}$$

To derive this conditional PC probability, we will first condition on $X_{r:n}=x_r$. This gives us:

$$\begin{aligned}
\pi_{2321} &= \mathbb{P}_{D,D^*}\Bigg\{d^*\Bigg[2-\left(\frac{1}{r}+\frac{1}{d^*}\right)\sum_{i=1}^{r-1}X_{i:n}-\left(\frac{n-r+1}{r}+\frac{d^*-r+1}{d^*}\right)x_r-\frac{(n-d^*)T^*}{d^*}\Bigg]\\
&\leq \sum_{i=r+1}^{d^*}(X_{i:n}-x_r)\\
&\leq d^*\Bigg[\left(\frac{1}{r}-\frac{1}{d^*}\right)\sum_{i=1}^{r-1}X_{i:n}+\left(\frac{n-r+1}{r}-\frac{d^*-r+1}{d^*}\right)x_r-\frac{(n-d^*)T^*}{d^*}\Bigg]\Bigg\}. \qquad (4.4.17)
\end{aligned}$$

Using similar steps for computing π_{2322} from Equation (4.4.13), we have that:

$$\pi_{2322} = \mathbb{P}_{D,D^*}\Bigg\{d^*\Bigg[\left(\frac{1}{r}-\frac{1}{d^*}\right)\sum_{i=1}^{r-1}X_{i:n}+\left(\frac{n-r+1}{r}-\frac{d^*-r+1}{d^*}\right)x_r-\frac{(n-d^*)T^*}{d^*}\Bigg]$$

$$\leq \sum_{i=r+1}^{d^*} (X_{i:n} - x_r)$$
$$\leq d^* \left[2 - \left(\frac{1}{r} + \frac{1}{d^*} \right) \sum_{i=1}^{r-1} X_{i:n} - \left(\frac{n-r+1}{r} + \frac{d^*-r+1}{d^*} \right) x_r - \frac{(n-d^*)T^*}{d^*} \right] \Bigg\}. \quad (4.4.18)$$

Now we focus on finding the distribution of the middle expression of the inequality found in Equations (4.4.17) and (4.4.18), which is $\sum_{i=r+1}^{d^*}(X_{i:n} - x_r)$. Notice that $X_{r+1:n} \leq X_{r+2:n} \leq \cdots \leq X_{d^*:n}$ are order statistics from a complete sample of size $(d^* - r)$ from an exponential distribution doubly truncated between $[x_r, T^*]$. Again, using a result from Lemma (2.1.4) Let

$$\sum_{i=r+1}^{d^*} (X_{i:n} - x_r) = \sum_{i=1}^{d^*-r} Y_{i:n-r},$$

where $Y_{1:n-r} \leq Y_{2:n-r} \leq \cdots \leq Y_{d^*-r:n-r}$ represent order statistics from a complete sample of size $(d^* - r)$ that is now only right-truncated at $T'' = T^* - x_r$. For notational convenience, $B_{23} := \sum_{i=1}^{d^*-r} Y_{i:n-r}$.

Using Proposition (3.2.1) where $m = d^* - r$ and $L = T''$, we propose the pdf for B_{23} to be:

$$f_{B_{23}}(b) = \frac{1}{(1 - e^{-T''})^{(d^*-r)}} \sum_{i=0}^{d^*-r} \binom{d^*-r}{i} (-1)^i e^{-T''i}$$
$$\times \frac{1}{\Gamma(d^*-r)} (b - T''i)^{(d^*-r)-1} e^{-(b-T''i)} I(b > T''i). \quad (4.4.19)$$

Furthermore, $0 \leq B_{23} \leq (d^* - r)T''$.

Next, we aim to find the pdf of $A_{11} := \sum_{i=1}^{r-1} X_{i:n}$. The corresponding result was established earlier in comparison 1; see Equation (3.2.17).

Finally, let us focus on finding the pdf for $X_{r:n}$. Since we have $D \in \{0, 1, \ldots, r-1\}$, we know that $X_{r:n} > T$. But since $D^* \in \{r+1, r+2, \ldots, n\}$ we have $X_{r:n} \leq T^*$. That is, conditioning on both $D = d, D^* = d^*$ we have that $X_{r:n}$ is the $(r-d)$th order statistic from $(d^* - d)$ observations truncated between $(T, T^*]$. Thus, using Lemma (2.1.2), we can derive the the pdf for $X_{r:n}$ conditional on $D = d$ and $D^* = d^*$:

$$
\begin{aligned}
&f_{X_{r:n}|D=d,D^*=d^*}(x_r)\\
&= \frac{(d^*-d)!}{((r-d)-1)!((d^*-d)-(r-d))!}\{F_{X_r}(x_r)\}^{(r-d)-1}\{1-F_{X_r}(x_r)\}^{(d^*-d)-(r-d)} f_{X_r}(x_r)\\
&= \frac{(d^*-d)!}{(r-d-1)!(d^*-r)!}\left(\frac{e^{-T}-e^{-x_r}}{e^{-T}-e^{-T^*}}\right)^{(r-d-1)}\left(1-\frac{e^{-T}-e^{-x_r}}{e^{-T}-e^{-T^*}}\right)^{(d^*-r)}\left(\frac{e^{-x_r}}{e^{-T}-e^{-T^*}}\right)\\
&= \frac{(d^*-d)!}{(r-d-1)!(d^*-r)!}\left(\frac{e^{-T}-e^{-x_r}}{e^{-T}-e^{-T^*}}\right)^{(r-d-1)}\left(\frac{e^{-x_r}-e^{-T^*}}{e^{-T}-e^{-T^*}}\right)^{(d^*-r)}\left(\frac{e^{-x_r}}{e^{-T}-e^{-T^*}}\right)\\
&= \frac{(d^*-d)!}{(r-d-1)!(d^*-r)!}\frac{1}{(e^{-T}-e^{-T^*})^{(d^*-d)}}(e^{-T}-e^{-x_r})^{(r-d-1)}(e^{-x_r}-e^{-T^*})^{(d^*-r)}(e^{-x_r}),\\
&\quad T < x_r < T^*.
\end{aligned} \tag{4.4.20}
$$

Now we want to return to the bounds expressed within π_{2321} and π_{2322}, which were shown in Equation (4.4.17) and (4.4.18). Consider the following:

$$
L_{232}(A_{11}, x_r, d^*) := d^*\left[2 - \left(\frac{1}{r} + \frac{1}{d^*}\right)A_{11} - \left(\frac{n-r+1}{r} + \frac{d^*-r+1}{d^*}\right)x_r - \frac{(n-d^*)T^*}{d^*}\right], \tag{4.4.21}
$$

$$
U_{232}(A_{11}, x_r, d^*) := d^*\left[\left(\frac{1}{r} - \frac{1}{d^*}\right)A_{11} + \left(\frac{n-r+1}{r} - \frac{d^*-r+1}{d^*}\right)x_r - \frac{(n-d^*)T^*}{d^*}\right]. \tag{4.4.22}
$$

Note that π_{2321} is associated with the case where $L_{232}(A_{11}, x_r, d^*) \leq U_{232}(A_{11}, x_r, d^*)$. Writing out this inequality, we obtain:

$$
\begin{aligned}
& L_{232}(A_{11}, x_r, d^*) \leq U_{232}(A_{11}, x_r, d^*) \\
& \Leftrightarrow 2 - \left(\frac{1}{r} + \frac{1}{d^*}\right) A_{11} - \left(\frac{n-r+1}{r} + \frac{d^*-r+1}{d^*}\right) x_r \\
& \leq \left(\frac{1}{r} - \frac{1}{d^*}\right) A_{11} + \left(\frac{n-r+1}{r} - \frac{d^*-r+1}{d^*}\right) x_r \\
& \Leftrightarrow 2 - A_{11}\left(\frac{1}{r} + \cancel{\frac{1}{d^*}} + \frac{1}{r} - \cancel{\frac{1}{d^*}}\right) \\
& \leq x_r\left(\frac{n-r+1}{r} - \cancel{\frac{d^*-r+1}{d^*}} + \frac{n-r+1}{r} + \cancel{\frac{d^*-r+1}{d^*}}\right) \\
& \Leftrightarrow 1 - \frac{A_{11}}{r} \leq \frac{x_r(n-r+1)}{r} \\
& \Leftrightarrow -A_{11} \leq x_r(n-r+1) - r \\
& \Leftrightarrow A_{11} \geq r - x_r(n-r+1).
\end{aligned} \tag{4.4.23}
$$

Similarly, $L_{232}(A_{11}, x_r, d^*) \geq U_{232}(A_{11}, x_r, d^*)$ implies that $A_{11} \leq r - x_r(n-r+1)$. Hence, utilising the total law of probability, we can compute the π_{2321} and π_{2322} using:

$$
\begin{aligned}
\pi_{2321} &= \int_T^{T^*} \int_{r-x_r(n-r+1)}^{(r-1)x_r} [F_{B_{23}}(U_{232}(a, x_r, d^*)) - F_{B_{23}}(L_{232}(a, x_r, d^*))] \\
& \times f_{X_{r:n}, A_{11}}(x_r, a)\, da\, dx_r \\
&= \int_T^{T^*} \int_{r-x_r(n-r+1)}^{(r-1)x_r} [F_{B_{23}}(U_{232}(a, x_r, d^*)) - F_{B_{23}}(L_{232}(a, x_r, d^*))] \\
& \times f_{A_{11}|X_r = x_r}(a) f_{X_{r:n}|D=d}(x_r)\, da\, dx_r,
\end{aligned} \tag{4.4.24}
$$

and for the other case, we have that

$$\begin{aligned}\pi_{2322} &= \int_T^{T^*}\int_0^{r-x_r(n-r+1)} [F_{B_{23}}(L_{232}(a,x_r,d^*)) - F_{B_{23}}(U_{232}(a,x_r,d^*))] \\ &\quad\times f_{X_{r:n},A_{11}}(x_r,a)\,da\,dx_r \\ &= \int_T^{T^*}\int_0^{r-x_r(n-r+1)} [F_{B_{23}}(L_{232}(a,x_r,d^*)) - F_{B_{23}}(U_{232}(a,x_r,d^*))] \\ &\quad\times f_{A_{11}|X_r=x_r}(a) f_{X_{r:n}|D=d}(x_r)\,da\,dx_r. \end{aligned} \tag{4.4.25}$$

Thus, the PC probability corresponding to the last case is:

$$\begin{aligned} &\sum_{d=0}^{r-1}\mathbb{P}(D=d)\sum_{d^*=r}^{n}\mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{233}-1|\le|\hat{\theta}_{231}-1|) \\ &= \sum_{d=0}^{r-1}\mathbb{P}(D=d)\left(\mathbb{P}_D(D^*=r)\pi_{231} + \sum_{d^*=r+1}^{n}\mathbb{P}_D(D^*=d^*)[\pi_{2321}+\pi_{2322}]\right) \\ &= \sum_{d=0}^{r-1}\mathbb{P}(D=d)\left(\mathbb{P}_D(D^{'}=r-d)\pi_{231} + \sum_{d'=r+1-d}^{n-d}\mathbb{P}_D(D'=d')[\pi_{2321}+\pi_{2322}]\right). \end{aligned} \tag{4.4.26}$$

4.5 Pitman Closeness Criterion for Comparison of $\hat{\theta}_1$ and $\hat{\theta}_3$

Having derived expressions for the PC probability in three distinct cases, we now combine the results to obtain the PC probability between the estimators $\hat{\theta}_1$ and $\hat{\theta}_3$. Specifically, we have:

$$\mathbb{P}(|\hat{\theta}_3-1|\le|\hat{\theta}_1-1|)$$

$$
\begin{aligned}
&= \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=d}^{r-1} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{213}-1| \leq |\hat{\theta}_{211}-1|) \\
&+ \sum_{d=r}^{n} \mathbb{P}(D=d) \sum_{d^*=d}^{n} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{223}-1| \leq |\hat{\theta}_{221}-1|) \\
&+ \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=r}^{n} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{233}-1| \leq |\hat{\theta}_{231}-1|) \\
&= \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d'=1}^{r-d-1} \mathbb{P}_D(D'=d') \\
&+ \sum_{d=r}^{n} \mathbb{P}(D=d)\left(\mathbb{P}_D(D'=0)\pi_{221} + \sum_{d'=1}^{n-d} \mathbb{P}_D(D'=d')[\pi_{2221}+\pi_{2222}]\right) \\
&+ \sum_{d=0}^{r-1} \mathbb{P}(D=d)\left(\mathbb{P}_D(D'=r-d)\pi_{231} + \sum_{d'=r+1-d}^{n-d} \mathbb{P}_D(D'=d')[\pi_{2321}+\pi_{2322}]\right), \qquad (4.5.1)
\end{aligned}
$$

where we have $\mathbb{P}(D = d)$ and $\mathbb{P}_D(D' = d')$ from Equations (3.1.7) and (4.1.9), respectively. The computational results are presented in Chapter 5.2, and the corresponding R code is provided in Appendix A.3.

Chapter 5

Numerical Results

This chapter presents the numerical results computing the Pitman closeness probabilities for different settings and with varying n, r, s, T, and T^*.

5.1 Comparison of $\hat{\theta}_1$ and $\hat{\theta}_2$

For the first comparison, we consider two possible values for n with varying values of r, s, and T, with the restriction that $r < s$. The case for $n = 10$ can be found in Table 5.1 and Figure 5.1, and $n = 15$ can be found in Table 5.2 and Figure 5.2. Each cell represents the result obtained by Equation (3.5.1). We see that the estimator based on s is always Pitman closer to θ than the estimator based on r, as expected. The code used for these results can be found in Appendix A.2.

Interestingly, as the number of observed failures associated with $\hat{\theta}_2$ increases (s), the PC probability decreases. However, the PC probability is still always above the 0.5 threshold. This may imply that for a fixed T, a substantial increase in the number of observed failures may not be necessary.

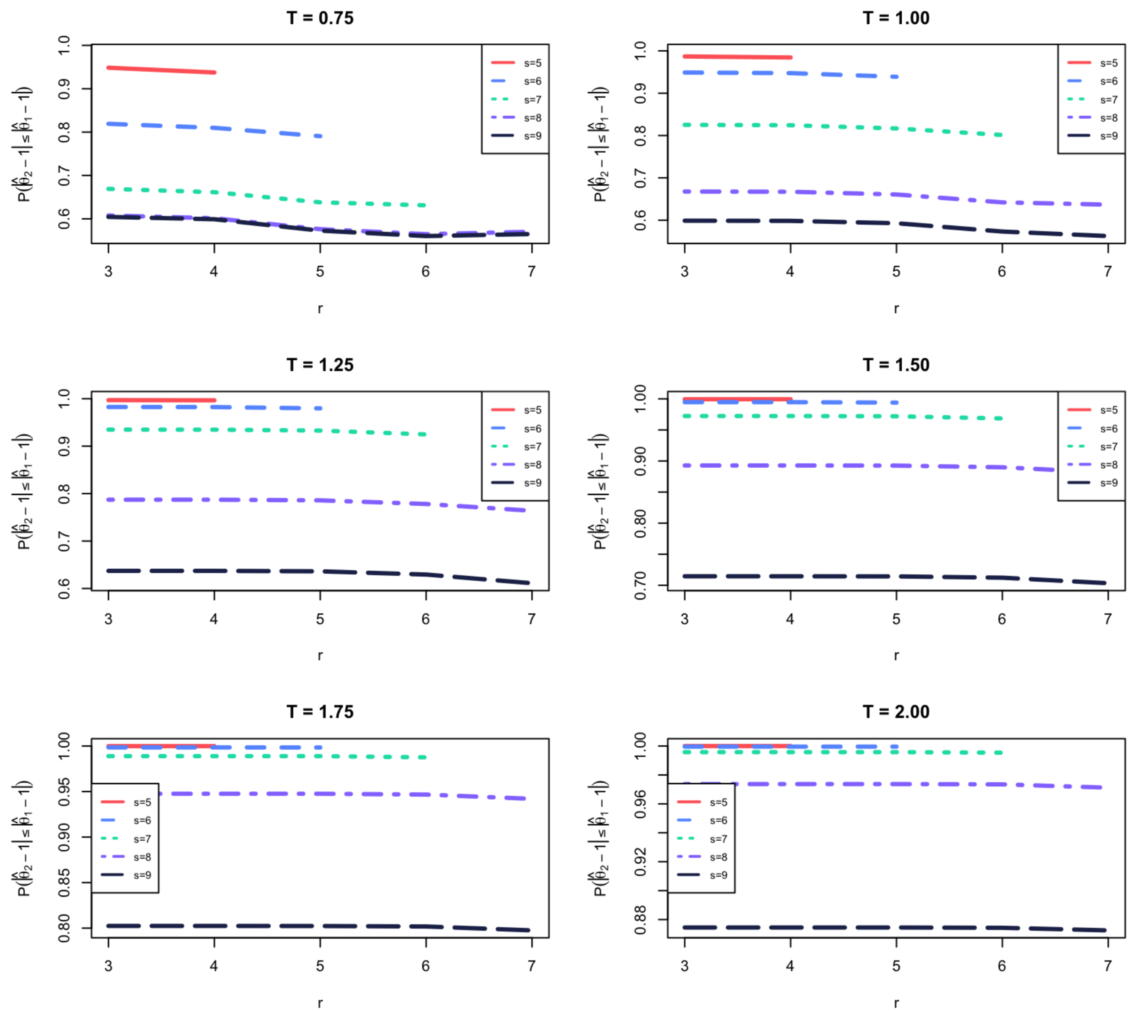

Figure 5.1: Line plots featuring the Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_2$ for $n = 10$ and varying values of $r, s,$ and T.

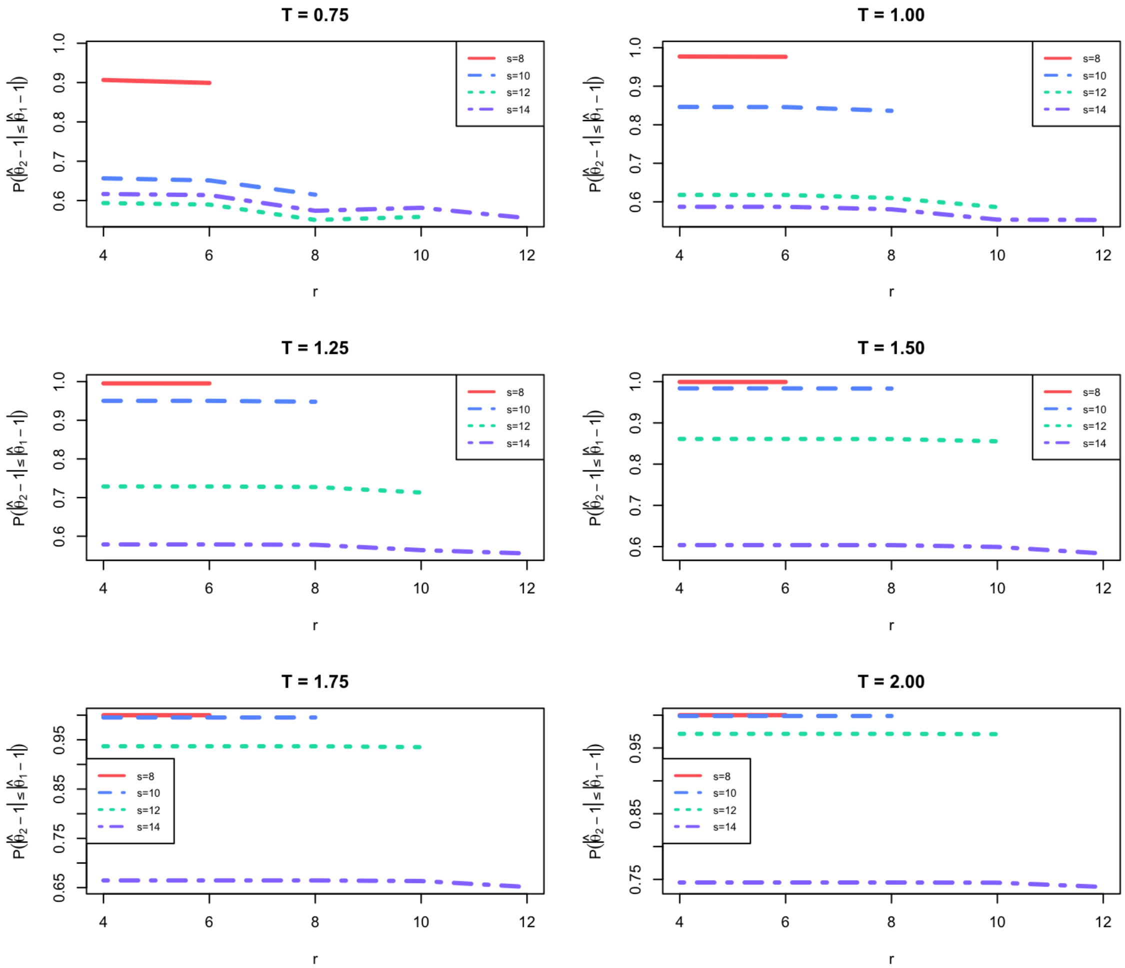

Figure 5.2: Line plots featuring the Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_2$ for $n = 15$ and varying values of $r, s,$ and T.

		T					
r	s	0.75	1.00	1.25	1.50	1.75	2.00
3	4	0.9854	0.998	0.9997	1.000	1.000	1.000
	5	0.9485	0.9868	0.9971	0.9994	0.9999	1.000
	6	0.8191	0.9489	0.9826	0.9947	0.9985	0.9996
	7	0.6691	0.8252	0.9349	0.9724	0.9889	0.9958
	8	0.6077	0.6677	0.7872	0.8928	0.9476	0.9737
	9	0.6044	0.5988	0.6372	0.7145	0.8024	0.8745
4	5	0.9375	0.9844	0.9967	0.9994	0.9999	1.000
	6	0.8098	0.9474	0.9824	0.9947	0.9985	0.9996
	7	0.6614	0.8243	0.9348	0.9724	0.9889	0.9958
	8	0.6013	0.6672	0.7872	0.8928	0.9476	0.9737
	9	0.5990	0.5984	0.6372	0.7145	0.8024	0.8745
5	6	0.7906	0.9388	0.9797	0.9940	0.9984	0.9996
	7	0.6380	0.8167	0.9329	0.9720	0.9889	0.9958
	8	0.5761	0.6607	0.7858	0.8926	0.9476	0.9737
	9	0.5727	0.5928	0.6362	0.7143	0.8023	0.8745
6	7	0.6311	0.8013	0.9248	0.9685	0.9876	0.9954
	8	0.5646	0.6421	0.7782	0.8899	0.9467	0.9735
	9	0.5602	0.5732	0.6294	0.7122	0.8018	0.8743
7	8	0.5702	0.6367	0.7638	0.8814	0.9420	0.9712
	9	0.5649	0.5625	0.6111	0.7033	0.7975	0.8725

Table 5.1: Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_2$ for $n = 10$ and varying values of r, s, and T.

		T					
r	s	0.75	1.00	1.25	1.50	1.75	2.00
4	6	0.9874	0.9987	0.9999	1.0000	1.0000	1.0000
	8	0.9066	0.9771	0.9955	0.9993	0.9999	1.0000
	10	0.6565	0.8463	0.9503	0.9837	0.9953	0.9988
	12	0.5936	0.6180	0.7287	0.8611	0.9369	0.9716
	14	0.6167	0.5872	0.5787	0.6036	0.6646	0.7453
6	8	0.8991	0.9763	0.9954	0.9993	0.9999	1.0000
	10	0.6512	0.8459	0.9503	0.9837	0.9953	0.9988
	12	0.5898	0.6179	0.7287	0.8611	0.9369	0.9716
	14	0.6139	0.5871	0.5787	0.6036	0.6646	0.7453
8	10	0.6150	0.8362	0.9480	0.9833	0.9953	0.9988
	12	0.5507	0.6099	0.7273	0.8609	0.9369	0.9716
	14	0.5741	0.5806	0.5778	0.6035	0.6646	0.7453
10	12	0.5586	0.5863	0.7131	0.8553	0.9351	0.9711
	14	0.5816	0.5539	0.5640	0.5991	0.6634	0.7450
12	14	0.5555	0.5530	0.5554	0.5833	0.6514	0.7385

Table 5.2: Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_2$ for $n = 15$ and varying values of r, s, and T.

It is worth noting that the PC probability also accounts for cases where the estimators are equal. The cases regarding ties are generally uninformative; if increasing the sample size or experiment duration produces the same estimator, there is no insight to be gained for estimating the true parameter. To illustrate how $\mathbb{P}(|\hat{\theta}_2 - 1| \leq |\hat{\theta}_1 - 1|)$ changes once we condition on the estimators being unequal, we provide additional tables with three entries for each combination of r, s, and T. In each cell (same combination of r, s, and T):

- The first row represents the conditional probability $\mathbb{P}(|\hat{\theta}_2 - 1| \leq |\hat{\theta}_1 - 1|)$ given that the estimators are different,
- the second row represents the conditional probability $\mathbb{P}(|\hat{\theta}_1 - 1| \leq |\hat{\theta}_2 - 1|)$ given that the estimators are different,
- and the last row represents the probability that the estimators tie; that is, $\mathbb{P}(|\hat{\theta}_1 - 1| = |\hat{\theta}_2 - 1|)$.

By construction, the probability that they tie is explicitly written for the third case where $D \in \{s, \ldots, n\}$ as outlined in the subchapter (3.4). As a clarification, this second set of tables cannot be used for PC comparisons since the equality case is included for the original definition and the criterion.

In Tables 5.3 and 5.4, we consider $n = 10$ and values $T = 0.75$, 1.00, 1.25, 1.50, 1.75, 2.00, $r = 3, 4, 5, 6, 7$, and values of s that are greater than r but up to 9. Meanwhile, in Tables 5.5 and 5.6 consider the same values for T, $r = 4, 6, 8, 10, 12$, and even values of s that are greater than r but up to 14.

Overall, for smaller values of s, the likelihood of ties decreases. This supports the earlier observation that while using a larger number of observed failures improves estimator performance, the benefit tapers off as s becomes too large.

Additionally, for fixed values of r and s, increasing T leads to a higher probability of ties, while also raising the probability that $\mathbb{P}(|\hat{\theta}_2 - 1| \leq |\hat{\theta}_1 - 1|)$ given that they are different.

5.2 Comparison of $\hat{\theta}_1$ and $\hat{\theta}_3$

For the second comparison, we again consider sample sizes $n = 10$ and $n = 15$, examining various combinations of r, T, and T^*. Results corresponding to $n = 10$ are presented in Table 5.7 and Figure 5.3, while those for $n = 15$ appear in Table 5.8 and Figure 5.4. Each cell in the tables reflects values computed using Equation (4.5.1). The code used to generate these results is provided in Appendix A.3.

As anticipated, the estimator based on T^* is always Pitman closer to θ than T, however, for a fixed r there is no discernible pattern that significantly increasing T^* will make $\hat{\theta}_3$ Pitman closer to θ compared to $\hat{\theta}_1$.

		T					
r	s	0.75	1.00	1.25	1.50	1.75	2.00
3	4	0.8883	0.9406	0.9685	0.9833	0.9911	0.9952
		0.1117	0.0594	0.0315	0.0167	0.0089	0.0048
		0.8697	0.9655	0.9920	0.9983	0.9996	0.9999
	5	0.8343	0.8880	0.9236	0.9481	0.9649	0.9763
		0.1657	0.1120	0.0764	0.0519	0.0351	0.0237
		0.6889	0.8824	0.9619	0.9888	0.9970	0.9992
	6	0.6732	0.8232	0.8639	0.8953	0.9195	0.9381
		0.3268	0.1768	0.1361	0.1047	0.0805	0.0619
		0.4465	0.7110	0.8719	0.9494	0.9816	0.9937
	7	0.5752	0.6729	0.7932	0.8326	0.8604	0.8834
		0.4248	0.3271	0.2068	0.1674	0.1396	0.1166
		0.2210	0.4656	0.6852	0.8350	0.9207	0.9642
	8	0.5750	0.5714	0.6334	0.7271	0.7859	0.8163
		0.4250	0.4286	0.3666	0.2729	0.2141	0.1837
		0.0770	0.2247	0.4196	0.6073	0.7551	0.8568
	9	0.5977	0.5688	0.5621	0.5861	0.6340	0.6868
		0.4023	0.4312	0.4379	0.4139	0.3660	0.3132
		0.0166	0.0695	0.1715	0.3101	0.4600	0.5992
4	5	0.7991	0.8672	0.9128	0.9428	0.9623	0.9751
		0.2009	0.1328	0.0872	0.0572	0.0377	0.0249
		0.6889	0.8824	0.9619	0.9888	0.9970	0.9992
	6	0.6564	0.8180	0.8625	0.8949	0.9194	0.9381
		0.3436	0.1820	0.1375	0.1051	0.0806	0.0619
		0.4465	0.7110	0.8719	0.9494	0.9816	0.9937
	7	0.5653	0.6712	0.7929	0.8326	0.8604	0.8834
		0.4347	0.3288	0.2071	0.1674	0.1396	0.1166
		0.2210	0.4656	0.6852	0.8350	0.9207	0.9642
	8	0.5681	0.5707	0.6334	0.7271	0.7859	0.8163
		0.4319	0.4293	0.3666	0.2729	0.2141	0.1837
		0.0770	0.2247	0.4196	0.6073	0.7551	0.8568
	9	0.5922	0.5684	0.5621	0.5861	0.6340	0.6868
		0.4078	0.4316	0.4379	0.4139	0.3660	0.3132
		0.0166	0.0695	0.1715	0.3101	0.4600	0.5992

Table 5.3: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_2 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_2 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_2 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 10$ across varying r, s, and T values (Part 1).

		T					
r	s	0.75	1.00	1.25	1.50	1.75	2.00
5	6	0.6217	0.7881	0.8412	0.8814	0.9113	0.9334
		0.3783	0.2119	0.1588	0.1186	0.0887	0.0666
		0.4465	0.7110	0.8719	0.9494	0.9816	0.9937
	7	0.5353	0.6569	0.7868	0.8303	0.8596	0.8831
		0.4647	0.3431	0.2132	0.1697	0.1404	0.1169
		0.2210	0.4656	0.6852	0.8350	0.9207	0.9642
	8	0.5408	0.5623	0.6310	0.7265	0.7858	0.8163
		0.4592	0.4377	0.3690	0.2735	0.2142	0.1837
		0.0770	0.2247	0.4196	0.6073	0.7551	0.8568
	9	0.5655	0.5624	0.5609	0.5860	0.6339	0.6868
		0.4345	0.4376	0.4391	0.4140	0.3661	0.3132
		0.0166	0.0695	0.1715	0.3101	0.4600	0.5992
6	7	0.5264	0.6281	0.7611	0.8094	0.8443	0.8726
		0.4736	0.3719	0.2389	0.1906	0.1557	0.1274
		0.2210	0.4656	0.6852	0.8350	0.9207	0.9642
	8	0.5283	0.5383	0.6179	0.7195	0.7824	0.8148
		0.4717	0.4617	0.3821	0.2805	0.2176	0.1852
		0.0770	0.2247	0.4196	0.6073	0.7551	0.8568
	9	0.5528	0.5413	0.5527	0.5828	0.6329	0.6865
		0.4472	0.4587	0.4473	0.4172	0.3671	0.3135
		0.0166	0.0695	0.1715	0.3101	0.4600	0.5992
7	8	0.5344	0.5314	0.5930	0.6979	0.7631	0.7989
		0.4656	0.4686	0.4070	0.3021	0.2369	0.2011
		0.0770	0.2247	0.4196	0.6073	0.7551	0.8568
	9	0.5575	0.5299	0.5307	0.5700	0.6250	0.6819
		0.4425	0.4701	0.4693	0.4300	0.3750	0.3181
		0.0166	0.0695	0.1715	0.3101	0.4600	0.5992

Table 5.4: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_2 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_2 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_2 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 10$ across varying r, s, and T values (Part 2).

		T					
r	s	0.75	1.00	1.25	1.50	1.75	2.00
4	6	0.8809	0.9292	0.9577	0.9748	0.9851	0.9912
		0.1191	0.0708	0.0423	0.0252	0.0149	0.0088
		0.8943	0.9818	0.9975	0.9997	1.0000	1.0000
	8	0.7744	0.8416	0.8836	0.9142	0.9368	0.9534
		0.2256	0.1584	0.1164	0.0858	0.0632	0.0466
		0.5862	0.8553	0.9614	0.9914	0.9983	0.9997
	10	0.5667	0.6889	0.7926	0.8299	0.8594	0.8833
		0.4333	0.3111	0.2074	0.1701	0.1406	0.1167
		0.2071	0.5058	0.7603	0.9043	0.9669	0.9897
	12	0.5815	0.5565	0.5903	0.6836	0.7538	0.7891
		0.4185	0.4435	0.4097	0.3164	0.2462	0.2109
		0.0291	0.1388	0.3377	0.5608	0.7439	0.8654
	14	0.6164	0.5830	0.5592	0.5494	0.5603	0.5905
		0.3836	0.4170	0.4408	0.4506	0.4397	0.4095
		0.0010	0.0100	0.0444	0.1203	0.2372	0.3780
6	8	0.7561	0.8359	0.8820	0.9138	0.9367	0.9534
		0.2439	0.1641	0.1180	0.0862	0.0633	0.0466
		0.5862	0.8553	0.9614	0.9914	0.9983	0.9997
	10	0.5600	0.6882	0.7925	0.8299	0.8594	0.8833
		0.4400	0.3118	0.2075	0.1701	0.1406	0.1167
		0.2071	0.5058	0.7603	0.9043	0.9669	0.9897
	12	0.5775	0.5563	0.5903	0.6836	0.7538	0.7891
		0.4225	0.4437	0.4097	0.3164	0.2462	0.2109
		0.0291	0.1388	0.3377	0.5608	0.7439	0.8654
	14	0.6136	0.5830	0.5592	0.5494	0.5603	0.5905
		0.3864	0.4170	0.4408	0.4506	0.4397	0.4095
		0.0010	0.0100	0.0444	0.1203	0.2372	0.3780

Table 5.5: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_2 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_2 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_2 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 15$ across varying r, s, and T values (Part 1).

		T					
r	s	0.75	1.00	1.25	1.50	1.75	2.00
8	10	0.5144	0.6684	0.7833	0.8260	0.8578	0.8827
		0.4856	0.3316	0.2167	0.1740	0.1422	0.1173
		0.2071	0.5058	0.7603	0.9043	0.9669	0.9897
	12	0.5372	0.5471	0.5883	0.6833	0.7537	0.7891
		0.4628	0.4529	0.4117	0.3167	0.2463	0.2109
		0.0291	0.1388	0.3377	0.5608	0.7439	0.8654
	14	0.5737	0.5764	0.5582	0.5493	0.5603	0.5905
		0.4263	0.4236	0.4418	0.4507	0.4397	0.4095
		0.0010	0.0100	0.0444	0.1203	0.2372	0.3780
10	12	0.5454	0.5196	0.5669	0.6706	0.7467	0.7854
		0.4546	0.4804	0.4331	0.3294	0.2533	0.2146
		0.0291	0.1388	0.3377	0.5608	0.7439	0.8654
	14	0.5812	0.5493	0.5438	0.5443	0.5587	0.5901
		0.4188	0.4507	0.4562	0.4557	0.4413	0.4099
		0.0010	0.0100	0.0444	0.1203	0.2372	0.3780
12	14	0.5551	0.5485	0.5347	0.5263	0.5430	0.5796
		0.4449	0.4515	0.4653	0.4737	0.4570	0.4204
		0.0010	0.0100	0.0444	0.1203	0.2372	0.3780

Table 5.6: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_2 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_2 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_2 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 15$ across varying r, s, and T values (Part 2).

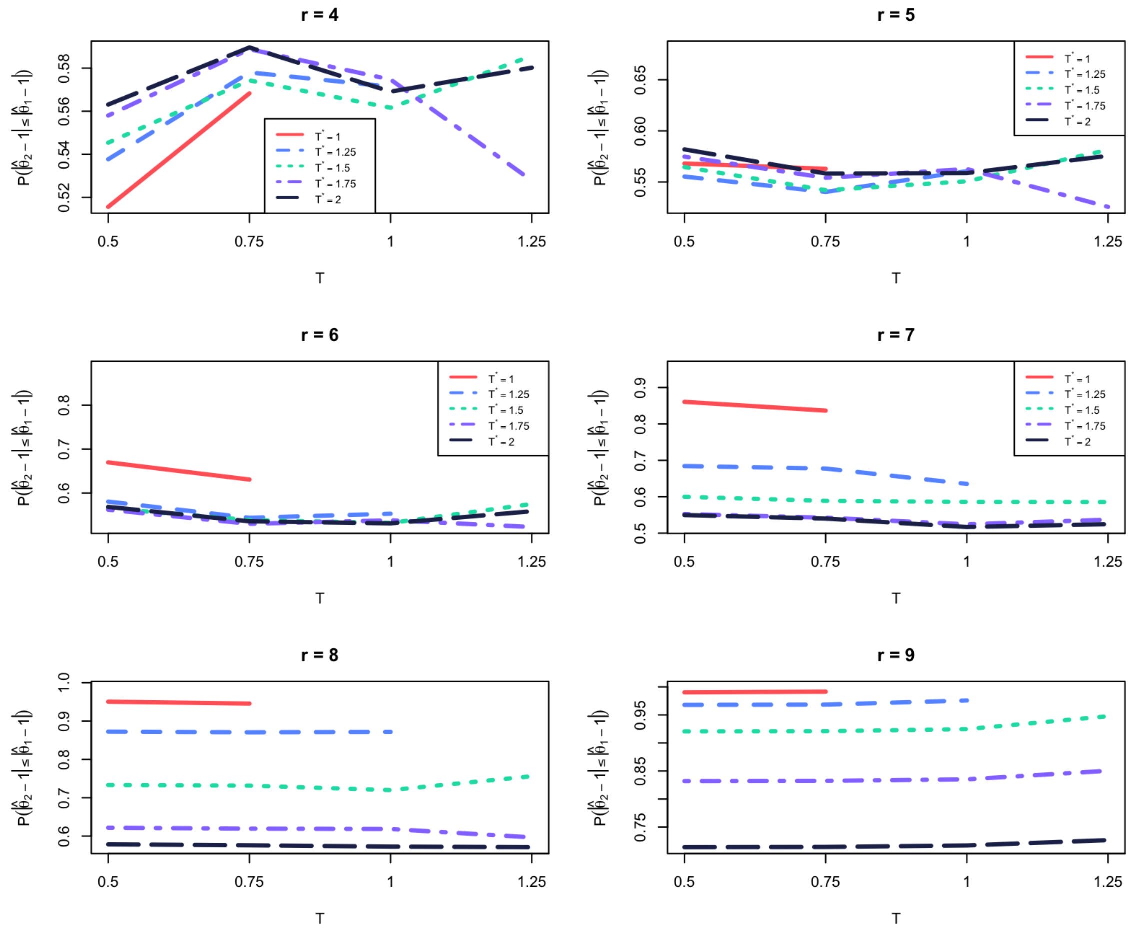

Figure 5.3: Line plots featuring the Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_3$ for $n = 10$ and varying values of r, T, and T^*.

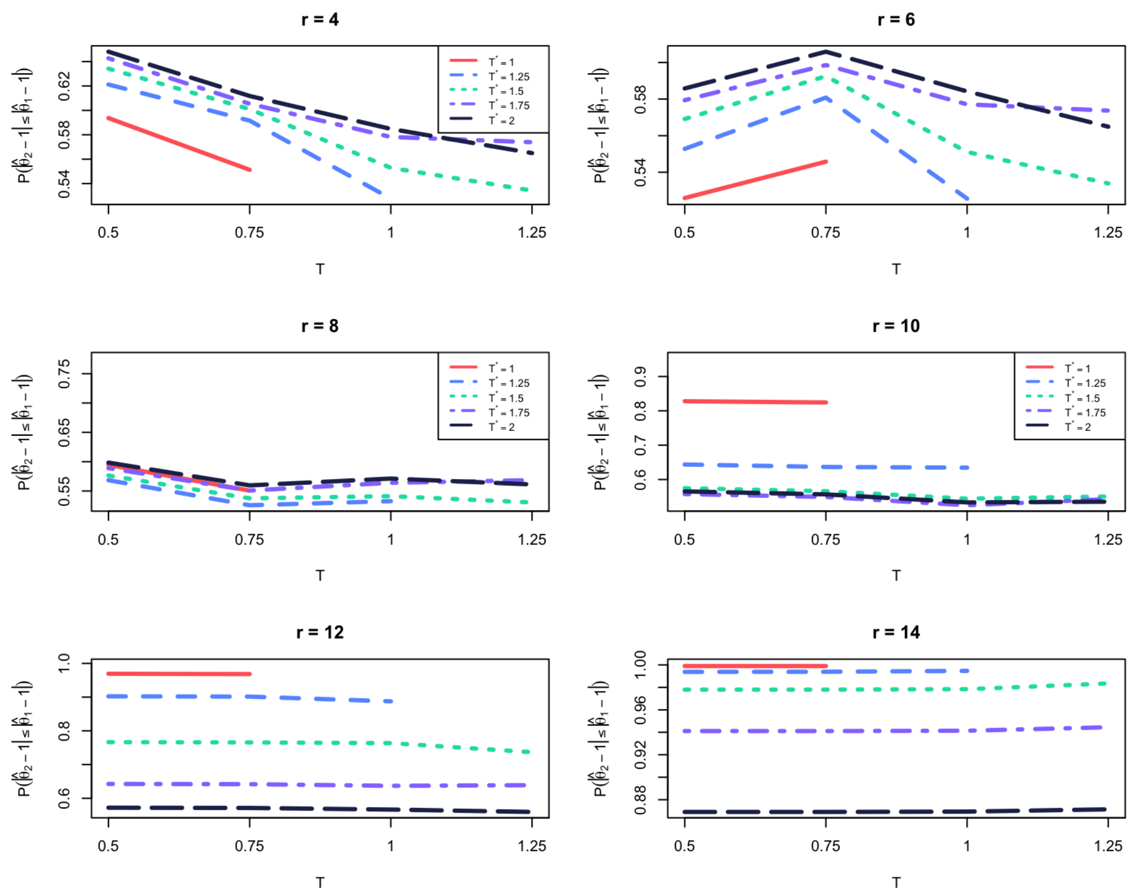

Figure 5.4: Line plots featuring the Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_3$ for $n = 15$ and varying values of r, T, and T^*.

		r					
T	T^*	4	5	6	7	8	9
0.50	0.75	0.5364	0.6816	0.8854	0.9546	0.9868	0.9984
	1.00	0.5156	0.5681	0.6698	0.8607	0.9506	0.9903
	1.25	0.5377	0.5553	0.5805	0.6843	0.8724	0.9679
	1.50	0.5454	0.5646	0.5652	0.5999	0.7332	0.9206
	1.75	0.5580	0.5748	0.5620	0.5525	0.6218	0.8320
	2.00	0.5631	0.5819	0.5685	0.5497	0.5785	0.7141
0.75	1.00	0.5683	0.5629	0.6307	0.8365	0.9456	0.9915
	1.25	0.5780	0.5403	0.5437	0.6774	0.8707	0.9684
	1.50	0.5743	0.5420	0.5386	0.5887	0.7316	0.9210
	1.75	0.5888	0.5541	0.5299	0.5424	0.6194	0.8324
	2.00	0.5896	0.5583	0.5357	0.5400	0.5760	0.7145
1.00	1.25	0.5714	0.5606	0.5529	0.6353	0.8718	0.9758
	1.50	0.5616	0.5507	0.5319	0.5859	0.7200	0.9247
	1.75	0.5747	0.5627	0.5380	0.5244	0.6185	0.8351
	2.00	0.5691	0.5587	0.5310	0.5169	0.5726	0.7172
1.25	1.50	0.5857	0.5814	0.5751	0.5856	0.7561	0.9477
	1.75	0.5278	0.5257	0.5227	0.5370	0.5967	0.8504
	2.00	0.5802	0.5756	0.5586	0.5247	0.5711	0.7267

Table 5.7: Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_3$ for $n = 10$ and varying values of r, T, and T^*.

		r					
T	T^*	4	6	8	10	12	14
0.50	0.75	0.6155	0.5662	0.7761	0.9528	0.9952	0.9999
	1.00	0.5937	0.5259	0.5945	0.8280	0.9694	0.9990
	1.25	0.6212	0.5528	0.5684	0.6437	0.9024	0.9938
	1.50	0.6342	0.5691	0.5764	0.5746	0.7665	0.9780
	1.75	0.6427	0.5794	0.5889	0.5575	0.6425	0.9411
	2.00	0.6481	0.5858	0.5984	0.5653	0.5721	0.8691
0.75	1.00	0.5513	0.5458	0.5506	0.8244	0.9684	0.9990
	1.25	0.5916	0.5808	0.5257	0.6366	0.9015	0.9939
	1.50	0.6010	0.5925	0.5375	0.5659	0.7656	0.9780
	1.75	0.6054	0.5986	0.5508	0.5492	0.6418	0.9411
	2.00	0.6118	0.6059	0.5596	0.5567	0.5713	0.8691
1.00	1.25	0.5277	0.5256	0.5327	0.6345	0.8875	0.9947
	1.50	0.5529	0.5511	0.5412	0.5442	0.7636	0.9784
	1.75	0.5782	0.5771	0.5640	0.5250	0.6369	0.9414
	2.00	0.5848	0.5842	0.5712	0.5333	0.5665	0.8694
1.25	1.50	0.5343	0.5339	0.5305	0.5505	0.7367	0.9835
	1.75	0.5739	0.5737	0.5685	0.5429	0.6388	0.9445
	2.00	0.5650	0.5649	0.5612	0.5354	0.5595	0.8714

Table 5.8: Pitman closeness probabilities between $\hat{\theta}_1$ and $\hat{\theta}_3$ for $n = 15$ and varying values of r, T, and T^*.

Similar to the previous section, we provide additional tables to show exactly how the probabilities $\mathbb{P}(|\hat{\theta}_3 - 1| \leq |\hat{\theta}_1 - 1|)$ and $\mathbb{P}(|\hat{\theta}_1 - 1| \leq |\hat{\theta}_3 - 1|)$ are affected once you account for the ties for each combination of r, T, and T^*. For each cell (same combination of r, T, and T^*):

- The first row represents the conditional probability $\mathbb{P}(|\hat{\theta}_3 - 1| \leq |\hat{\theta}_1 - 1|)$ given that the estimators are different,
- the second row represents the conditional probability $\mathbb{P}(|\hat{\theta}_1 - 1| \leq |\hat{\theta}_3 - 1|)$ given that the estimators are different,
- and the last row represents the probability that the estimators tie; that is, $\mathbb{P}(|\hat{\theta}_1 - 1| = |\hat{\theta}_3 - 1|)$.

To emphasise, once we condition on the estimators being different, these table no longer represent the Pitman closeness probabilities, but they represent a dissection after removing tie cases.

By construction, we know that the estimators tie in the first case where $D, D^* \in \{0, 1, \ldots, r-1\}$ as expressed in subchapter (4.2). However, there is another case where ties could occur. Recall that in the second case, we considered the scenario where $D, D^* \in \{r, r+1, \ldots, n\}$ and the estimators corresponding to this case are:

$$\hat{\theta}_{221} := \frac{1}{D}\left(\sum_{i=1}^{D} X_{i:n} + (n-D)T\right), \quad \hat{\theta}_{223} := \frac{1}{D^*}\left(\sum_{i=1}^{D^*} X_{i:n} + (n-D^*)T^*\right). \tag{5.2.1}$$

If $D = D^* = n$, then the estimators are equal. In fact, we have a complete sample. Hence, consider the following decomposition instead:

$$\mathbb{P}(|\hat{\theta}_3 - 1| \leq |\hat{\theta}_1 - 1|)$$

$$
\begin{aligned}
&= \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=d}^{r-1} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{213}-1| \leq |\hat{\theta}_{211}-1|) \\
&+ \sum_{d=r}^{n} \mathbb{P}(D=d) \sum_{d^*=d}^{n} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{223}-1| \leq |\hat{\theta}_{221}-1|) \\
&+ \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=r}^{n} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{233}-1| \leq |\hat{\theta}_{231}-1|) \\
&= \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=d}^{r-1} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{213}-1| \leq |\hat{\theta}_{211}-1|) \\
&+ \mathbb{P}(D=n)\mathbb{P}_D(D^*=n)\mathbb{P}_{D,D^*}(|\hat{\theta}_{223}-1| \leq |\hat{\theta}_{221}-1|) \\
&+ \sum_{d=r}^{n-1} \mathbb{P}(D=d) \sum_{d^*=d}^{n} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{223}-1| \leq |\hat{\theta}_{221}-1|) \\
&+ \sum_{d=0}^{r-1} \mathbb{P}(D=d) \sum_{d^*=r}^{n} \mathbb{P}_D(D^*=d^*)\mathbb{P}_{D,D^*}(|\hat{\theta}_{233}-1| \leq |\hat{\theta}_{231}-1|),
\end{aligned}
\tag{5.2.2}
$$

where the first two terms of the last equality correspond to the probability of ties.

Tables 5.9 and 5.10 feature cases where $n = 10$, $r = 4, 5, 6, 7, 8, 9$, $T = 0.50$, 0.75, 1.00, 1.25, and values of T^* that are greater than T but up to 2.00. Furthermore, in Tables 5.11 and 5.12 we consider same values for T and T^* as above, but now $r = 4, 6, 8, 10, 12, 14$.

Although Table 5.7 and Table 5.8 suggest that $\hat{\theta}_3$ is always Pitman closer to θ than $\hat{\theta}_1$, somewhat unexpectedly, there are instances where $\mathbb{P}(|\hat{\theta}_1 - 1| < |\hat{\theta}_3 - 1|)$ exceeds 0.5. We highlight such cases in bold text in the tables.

Nonetheless, $\mathbb{P}(|\hat{\theta}_1 - 1| < |\hat{\theta}_3 - 1|)$ does not reflect the actual PC probabilities and therefore the PC criterion cannot be applied to this case. However, these tables indicate that the apparent advantage of $\hat{\theta}_3$ from the Pitman closeness criterion may largely be due to the probability of ties. This is similar to the results mentioned by Davies [13], where similar estimators under a Type-I HCS, revealed instances where conditional on the estimators being different, there is a higher probability that the estimator based on the shorter termination time T is closer to θ compared to the estimator based on the termination time based on T^*.

		r					
T	T^*	4	5	6	7	8	9
0.50	0.75	**0.4669**	0.5378	0.7434	0.7944	0.8288	0.9048
		0.5331	0.4622	0.2566	0.2056	0.1712	0.0952
		0.1304	0.3112	0.5535	0.7791	0.9231	0.9834
	1.00	**0.4982**	0.5104	0.5355	0.7008	0.7802	0.8608
		0.5018	0.4896	0.4645	0.2992	0.2198	0.1392
		0.0346	0.1177	0.2891	0.5344	0.7754	0.9306
	1.25	0.5339	0.5377	0.5188	0.5392	0.6959	0.8130
		0.4661	0.4623	0.4812	0.4608	0.3041	0.1870
		0.0081	0.0382	0.1282	0.3149	0.5805	0.8286
	1.50	0.5446	0.5596	0.5420	0.5207	0.5606	0.7437
		0.4554	0.4404	0.4580	0.4793	0.4394	0.2563
		0.0018	0.0113	0.0507	0.1651	0.3928	0.6900
	1.75	0.5578	0.5734	0.5537	0.5138	**0.4991**	0.6347
		0.4422	0.4266	0.4463	0.4862	0.5009	0.3653
		0.0004	0.0031	0.0185	0.0794	0.2449	0.5400
	2.00	0.5630	0.5815	0.5657	0.5330	0.5080	0.5228
		0.4370	0.4185	0.4343	0.4670	0.4920	0.4772
		0.0002	0.0009	0.0064	0.0358	0.1433	0.4009
0.75	1.00	0.5521	0.5037	**0.4795**	0.6477	0.7560	0.8744
		0.4479	0.4963	0.5205	0.3523	0.2440	0.1256
		0.0362	0.1193	0.2906	0.5360	0.7770	0.9322
	1.25	0.5738	0.5212	**0.4757**	0.5281	0.6905	0.8139
		0.4262	0.4788	0.5243	0.4719	0.3095	0.1861
		0.0097	0.0398	0.1298	0.3164	0.5821	0.8302
	1.50	0.5728	0.5360	0.5132	0.5065	0.5568	0.7439
		0.4272	0.4640	0.4868	0.4935	0.4432	0.2561
		0.0034	0.0128	0.0523	0.1667	0.3944	0.6916
	1.75	0.5880	0.5520	0.5203	0.5020	**0.4949**	0.6344
		0.4120	0.4480	0.4797	0.4980	0.5051	0.3656
		0.0020	0.0047	0.0201	0.0810	0.2465	0.5416
	2.00	0.5889	0.5572	0.5320	0.5221	0.5041	0.5223
		0.4111	0.4428	0.4680	0.4779	0.4959	0.4777
		0.0017	0.0025	0.0080	0.0374	0.1449	0.4024

Table 5.9: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_3 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_3 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_3 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 10$ across varying r, s, and T values (Part 1).

T	T^*	r = 4	5	6	7	8	9
1.00	1.25	0.5634	0.5383	**0.4812**	**0.4598**	0.6868	0.8503
		0.4366	0.4617	0.5188	0.5402	0.3132	0.1497
		0.0182	0.0483	0.1383	0.3250	0.5906	0.8387
	1.50	0.5563	0.5409	0.5016	**0.4979**	0.5310	0.7489
		0.4437	0.4591	0.4984	0.5021	0.4690	0.2511
		0.0119	0.0214	0.0608	0.1752	0.4029	0.7001
	1.75	0.5701	0.5568	0.5244	**0.4777**	**0.4879**	0.6334
		0.4299	0.4432	0.4756	0.5223	0.5121	0.3666
		0.0105	0.0132	0.0286	0.0895	0.2550	0.5501
	2.00	0.5647	0.5538	0.5232	**0.4936**	**0.4951**	0.5199
		0.4353	0.4462	0.4768	0.5064	0.5049	0.4801
		0.0103	0.0110	0.0165	0.0459	0.1534	0.4110
1.25	1.50	0.5702	0.5615	0.5357	**0.4825**	0.5745	0.8106
		0.4298	0.4385	0.4643	0.5175	0.4255	0.1894
		0.0359	0.0454	0.0848	0.1992	0.4269	0.7241
	1.75	0.5109	0.5074	**0.4962**	**0.4777**	**0.4407**	0.6486
		0.4891	0.4926	0.5038	0.5223	0.5593	0.3514
		0.0345	0.0372	0.0526	0.1135	0.2790	0.5741
	2.00	0.5653	0.5603	0.5399	**0.4889**	**0.4786**	0.5164
		0.4347	0.4397	0.4601	0.5111	0.5214	0.4836
		0.0343	0.0350	0.0405	0.0700	0.1774	0.4350

Table 5.10: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_3 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_3 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_3 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 10$ across varying r, s, and T values (Part 2).

		r					
T	T^*	4	6	8	10	12	14
0.50	0.75	0.6116	0.5150	0.6181	0.7722	0.8357	0.9316
		0.3884	0.4850	0.3819	0.2278	0.1643	0.0684
		0.0102	0.1057	0.4138	0.7929	0.9709	0.9990
	1.00	0.5934	0.5171	0.5259	0.6599	0.7793	0.8989
		0.4066	0.4829	0.4741	0.3401	0.2207	0.1011
		0.0008	0.0182	0.1447	0.4942	0.8612	0.9900
	1.25	0.6212	0.5517	0.5511	0.5314	0.7109	0.8613
		0.3788	0.4483	0.4489	0.4686	0.2891	0.1387
		0.0001	0.0025	0.0386	0.2397	0.6623	0.9556
	1.50	0.6342	0.5690	0.5727	0.5295	0.5836	0.8171
		0.3658	0.4310	0.4273	0.4705	0.4164	0.1829
		0.0000	0.0003	0.0086	0.0957	0.4392	0.8797
	1.75	0.6427	0.5794	0.5882	0.5424	0.5195	0.7514
		0.3573	0.4206	0.4118	0.4576	0.4805	0.2486
		0.0000	0.0000	0.0017	0.0331	0.2561	0.7628
	2.00	0.6481	0.5858	0.5983	0.5608	0.5055	0.6536
		0.3519	0.4142	0.4017	0.4392	0.4945	0.3464
		0.0000	0.0000	0.0003	0.0103	0.1346	0.6220
0.75	1.00	0.5509	0.5373	**0.4746**	0.6528	0.7725	0.9034
		0.4491	0.4627	0.5254	0.3472	0.2275	0.0966
		0.0009	0.0183	0.1447	0.4942	0.8613	0.9901
	1.25	0.5916	0.5798	0.5066	0.5219	0.7082	0.8614
		0.4084	0.4202	0.4934	0.4781	0.2918	0.1386
		0.0001	0.0026	0.0387	0.2398	0.6624	0.9557
	1.50	0.6010	0.5923	0.5334	0.5199	0.5821	0.8171
		0.3990	0.4077	0.4666	0.4801	0.4179	0.1829
		0.0001	0.0004	0.0086	0.0958	0.4393	0.8798
	1.75	0.6053	0.5985	0.5500	0.5337	0.5184	0.7515
		0.3947	0.4015	0.4500	0.4663	0.4816	0.2485
		0.0001	0.0001	0.0017	0.0332	0.2562	0.7629
	2.00	0.6118	0.6058	0.5594	0.5521	0.5046	0.6536
		0.3882	0.3942	0.4406	0.4479	0.4954	0.3464
		0.0001	0.0001	0.0004	0.0103	0.1347	0.6221

Table 5.11: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_3 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_3 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_3 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 15$ across varying r, s, and T values (Part 1).

		r					
T	T^*	4	6	8	10	12	14
1.00	1.25	0.5272	0.5239	0.5134	0.5186	0.6657	0.8779
		0.4728	0.4761	0.4866	0.4814	0.3343	0.1221
		0.0011	0.0035	0.0396	0.2408	0.6633	0.9566
	1.50	0.5524	0.5505	0.5367	**0.4954**	0.5777	0.8188
		0.4476	0.4495	0.4633	0.5046	0.4223	0.1812
		0.0010	0.0013	0.0096	0.0968	0.4402	0.8808
	1.75	0.5778	0.5767	0.5628	0.5082	0.5112	0.7517
		0.4222	0.4233	0.4372	0.4918	0.4888	0.2483
		0.0010	0.0011	0.0027	0.0341	0.2571	0.7639
	2.00	0.5844	0.5837	0.5707	0.5280	**0.4985**	0.6535
		0.4156	0.4163	0.4293	0.4720	0.5015	0.3465
		0.0010	0.0010	0.0013	0.0113	0.1356	0.6230
1.25	1.50	0.5314	0.5308	0.5234	**0.4994**	0.5252	0.8549
		0.4686	0.4692	0.4766	0.5006	0.4748	0.1451
		0.0063	0.0066	0.0149	0.1021	0.4455	0.8860
	1.75	0.5712	0.5710	0.5650	0.5242	0.5102	0.7596
		0.4288	0.4290	0.4350	0.4758	0.4898	0.2404
		0.0063	0.0064	0.0080	0.0394	0.2624	0.7692
	2.00	0.5622	0.5621	0.5582	0.5276	**0.4873**	0.6541
		0.4378	0.4379	0.4418	0.4724	0.5127	0.3459
		0.0063	0.0063	0.0066	0.0166	0.1409	0.6283

Table 5.12: Tables displaying the following probabilities: $\mathbb{P}(|\hat{\theta}_3 - \theta| < |\hat{\theta}_1 - \theta|)$, $\mathbb{P}(|\hat{\theta}_1 - \theta| < |\hat{\theta}_3 - \theta|)$, and $\mathbb{P}(|\hat{\theta}_3 - \theta| = |\hat{\theta}_1 - \theta|)$ with $n = 15$ across varying r, s, and T values (Part 2).

Chapter 6

Conclusion

Assuming lifetimes follow an exponential distribution with scale parameter θ, this thesis investigates the Pitman closeness (PC) criterion among maximum likelihood estimators (MLEs) of θ derived from observations drawn from a Type-II hybrid censored scheme (HCS). In this scheme, the lifetime experiment is terminated at $\max\{X_{r:n}, T\}$, where r denotes the number of observed failures and T is a pre-fixed termination time. The MLE of θ was initially derived by Childs et al. [11]. To derive the PC probabilities, classical results from order statistics [3] are employed, as well as utilising similar techniques from earlier work that derived the PC probability of MLEs of θ under Type-I hybrid censoring [13].

The first comparison in this thesis considers MLEs of θ based on the termination times $\max\{X_{r:n}, T\}$ and $\max\{X_{s:n}, T\}$ with $s > r$. The numerical results indicate that the estimator with respect to the second termination time is consistently Pitman closer to θ, suggesting that, for fixed T, observing more failures improves estimator performance.

Meanwhile, the second comparison examines MLEs of θ based on $\max\{X_{r:n}, T\}$

and $\max\{X_{r:n}, T^*\}$ with $T^* > T$. Here, the numerical results also reveal that the estimator with the latter termination time is Pitman closer to θ. However, in some cases, conditional on the estimators being different, there is a higher probability that the estimator based on the first termination time T is closer to θ compared to the estimator based on the termination time based on T^*.

These reveal that although extending the termination time leads to a better estimator under the Pitman closeness criterion, in large part this is because they tie often. Furthermore, large increases in the number of observed failures or experiment duration yield diminishing returns in PC, suggesting that slight extensions to the study time or observations may be adequate.

This thesis has several limitations. The exact probabilities were shown for specific values of n, r, s, T, and T^*; it is plausible that other values would reveal different patterns. For instance, cases might exist where the MLE based on the termination time associated with a larger number of observed failures is not necessarily Pitman closer to θ than the estimator associated with the termination time that focuses on a smaller number of observed failures. Additionally, the second comparison is more computationally demanding than the first due to the subcases used to derive the exact expressions, but there may be more efficient derivations for the PC probabilities.

Furthermore, Type-I and Type-II hybrid censoring schemes have inherent limitations. In Type-I HCS, no failures may occur before the pre-specified time T, whereas in Type-II HCS, T may exceed the time for $X_{n:n}$, the failure time of the last observation. These shortcomings motivate the introduction of generalised hybrid censoring schemes, which incorporate additional constraints to address these limitations of their non-generalised counterparts.

The Generalised Type-I HCS guarantees at least k observed lifetimes, preventing the case of having no failures. Meanwhile, the Generalised Type-II HCS incorporates an additional pre-specified termination time, allowing the experiment to conclude earlier. Specific details can be found in Balakrishnan et al [9].

Assuming the data arises from the exponential distribution, future work may consider comparing alternative estimators that do not arise from the MLE of θ under Type-II HCS to assess viable substitutes, or compare MLEs of θ under the Generalised Type-I HCS and the Generalised Type-II HCS.

Appendix A

R Code

A.1 Commonly Used Probability Density Functions

```
# Define f_A_{11}

f_A11 <- Vectorize(function(a, x_r, r){

  if(a > ((r-1)*x_r)){return(0)}

  total <- 0

  denom <- (1 - exp(-x_r))^(r-1)

  for(i in 0:(r-1)){

    part1 <- choose(r-1, i) * exp(-i*x_r) * (-1)^(i)

    q <- a - i*x_r

    part2 <- if((q > 0) && (a < (r-1)*x_r)){

      dgamma(q, shape = r - 1)

    } else {0}

    total <- total + part1*part2

  }
```

```
  return(total/denom)
}, vectorize.args = "a")

# Define f_A_{12}
f_A12 <- Vectorize(function(a, term1, d){
  total <- 0
  denom <- (1 - exp(-term1))^(d)
  for(i in 0:d){
    part1 <- choose(d, i) * exp(-i*term1) * (-1)^(i)
    q <- a - i*term1
    part2 <- if((q > 0) && (a < d*term1)){
      dgamma(q, shape = d)
    } else {0}
    total <- total + part1*part2
  }
  return(total/denom)
}, vectorize.args = "a")
```

A.2 Comparison of $\hat{\theta}_1$ and $\hat{\theta}_2$

We include the R code used to compute Equation (3.5.1) and results are shown in Section (5.1).

A.2.1 Comparison 1 Case 1

```
# Define f_{X_{r:n}}
comp1case1_fxrn <- Vectorize(function(x_r, n, r, term1, d){
  if(x_r > term1){
    part1 <- factorial(n-d)/(factorial(r-d-1) * factorial(n-r))
    ex <- exp(-x_r)/exp(-term1)
    part2 <- (1 - ex)^(r-d-1)
    part3 <- ex^(n-r+1)
    return(part1*part2*part3)
  } else {
    return(0)
  }
}, vectorize.args = "x_r")

L_11 <- function(a, x_r, n, r, s){
  part1 <- 2 - (1/r + 1/s) * a
  part2 <- ((n-r)/r + (n-r)/s + 1/r + 1/s) * x_r
  final <- s * (part1 - part2)
  return(final)
}

U_11 <- function(a, x_r, n, r, s){
  part1 <- (1/r - 1/s) * a
  part2 <- ((n-r)/r - (n-r)/s + 1/r -1/s) * x_r
```

```
  final <- s * (part1 + part2)
  return(final)
}

# The inner integrand
comp1case1_inner_int <- function(a, x_r, n, r, s, type = "111"){
  a1 <- L_11(a, x_r, n, r, s)
  a2 <- U_11(a, x_r, n, r, s)

  f1 <- pgamma(a1, shape = s-r, scale = 1)
  f2 <- pgamma(a2, shape = s-r, scale = 1)

  f_A_val <- f_A11(a, x_r, r)
  if(type == "111"){
    return((f2 - f1) * f_A_val)
  } else if (type == "112"){
    return((f1 - f2) * f_A_val)
  }
}

# The outer integrand for pi_111 and pi_112
comp1case1_outer_int <- Vectorize(function(x_r, n, r, s,
term1, d, type = "111"){
  if(type == "111"){
```

```
    lower_a <- r - x_r*(n-r+1)
    lower_a[lower_a < 0] <- 0
    upper_a <- (r-1)*x_r
    fit_call <- quote(function(a)
      comp1case1_inner_int(a, x_r, n, r, s, type = "111"))
  } else if (type == "112"){
    upper_a <- r - x_r*(n-r+1)
    lower_a <- 0
    fit_call <- quote(function(a)
      comp1case1_inner_int(a, x_r, n, r, s, type = "112"))
  }
  if (lower_a < upper_a) {
    inner_result <- integrate(eval(fit_call),
                    lower = lower_a, upper = upper_a)$value
  } else {
    inner_result <- 0
  }
  mult2 <- comp1case1_fxrn(x_r, n, r, term1, d)
  result <- inner_result * mult2
  return(result)
}, vectorize.args = "x_r")

# Comparison 1 Case 1 Full Computation
comp1case1 <- function(n, r, s, term1){
```

```
  prob_d <- 1 - exp(-term1)
  total <- 0
  for(d in 0:(r-1)){
    pid1 <- dbinom(d, size = n, prob = prob_d)
    pi_111 <- integrate(
        function(x_r) comp1case1_outer_int(x_r, n, r, s,
        term1, d, type = "111"),
        lower = term1, upper = Inf)$value
    pi_112 <- integrate(
        function(x_r) comp1case1_outer_int(x_r, n, r, s,
        term1, d, type = "112"),
        lower = term1, upper = Inf)$value
    total <- total + (pid1 * (pi_111 + pi_112))
  }
  return(total)
}
```

A.2.2 Comparison 1 Case 2

```
# Bounds for f_B_{12}
L_12 <- function(a, n, s, term1, d){
  theta_121 <- (a+(n-d)*term1)/d
  val <- 2*s - (theta_121 * (s+d))
  return(val)
}
```

```
U_12 <- function(a, n, s, term1, d){
  theta_121 <- (a+(n-d)*term1)/d
  val <- theta_121 * (s-d)
  return(val)
}

# General integrand computation (for pi_121 or pi_122)
comp1case22_integrand <- function(a, n, s, term1, d, type = "121"){
  a_121 <- L_12(a, n, s, term1, d)
  a_122 <- U_12(a, n, s, term1, d)

  f1 <- pgamma(a_121, shape = s-d, scale = 1)
  f2 <- pgamma(a_122, shape = s-d, scale = 1)
  f_A_val <- f_A12(a, term1, d)
  if(type == "121"){
    return((f2 - f1)*f_A_val)
  } else if (type == "122"){
    return((f1 - f2)*f_A_val)
  }
}

# Used to define pi_121 or pi_122
pi_12 <- function(n, s, term1, d, type = "121"){
```

```
  if(type == "121"){
    low_bd <- d - (n-d)*term1
    upp_bd <- d*term1
    fit_call <- quote(function(a)
      comp1case22_integrand(a, n, s, term1, d, "121"))
  } else if (type == "122"){
    low_bd <- 0
    upp_bd <- d - (n-d)*term1
    fit_call <- quote(function(a)
      comp1case22_integrand(a, n, s, term1, d, "122"))
  }
  if(low_bd < upp_bd){
    result = integrate(eval(fit_call),
        lower = low_bd, upper = upp_bd)$value
  } else {
    result = 0
  }
  return(result)
}

# Comparison 1 Case 2 Full Computation
comp1case2 = function(n, r, s, term1){
  total <- 0
  for(d in r:(s-1)){
```

```
    probd <- 1 - exp(-term1)
    PD <- dbinom(d, size = n, prob = probd)
    res1 <- pi_12(n, s, term1, d, type = "121")
    res2 <- pi_12(n, s, term1, d, type = "122")
    total <- total + (PD*(res1+res2))
  }
  return(total)
}
```

A.2.3 Comparison 1 Case 3

```
comp1case3 <- function(n, s, term1){
  probd <- 1 - exp(-term1)
  val <- 1 - pbinom(s - 1, size = n, prob = probd)
  return(val)
}
```

A.3 Comparison of $\hat{\theta}_1$ and $\hat{\theta}_3$

The R code used to compute Equation (4.5.1) is included here, and the corresponding results are presented in Section (5.2).

A.3.1 Comparison 2 Case 1

```
comp2case1 <- function(n, r, term1, term2){
  total <- 0
```

```
  for(d in 0:(r-1)){
    prob_d <- 1 - exp(-term1)
    pid1 <- dbinom(d, size = n, prob = prob_d)
    total2 <- 0
    for(dstar in d:(r-1)){
      dprime <- dstar - d
      prob_dprime <- 1 - exp(-(term2 - term1))
      pid2 <- dbinom(dprime, size = (n-d), prob = prob_dprime)
      total2 <- total2 + pid2
    }
    total <- total + pid1*total2
  }
  return(total)
}
```

A.3.2 Comparison 2 Case 2

```
U_221 <- function(n, term1, term2, d){
  m <- (n-d) * (term2-term1) / d
  part1 <- d * (1 - (m/2))
  part2 <- (n-d) * term1
  return(part1 - part2)
}


pi_2212 <- function(n, term1, term2, d){
```

```
  lower_bd <- 0
  upper_bd <- U_221(n, term1, term2, d)
  if(lower_bd <= upper_bd){
    result <- integrate(
        function(a) f_A12(a, term1, d), lower = lower_bd,
        upper = upper_bd)$value
  } else {
    result <- 0
  }
  return(result)
}

# Define f_{B_{22}}
f_B22 <- Vectorize(function(b, n, term1, term2, d, dstar){
  tprime <- term2 - term1
  dprime <- dstar - d
  total <- 0
  denom <- (1 - exp(-tprime))^dprime
  for(i in 0:dprime){
    part1 <- choose(dprime, i) * exp(-i*tprime) * (-1)^(i)
    q <- b - i*tprime
    part2 <- if((q > 0) && (b < (dprime*tprime))){
      dgamma(q, shape = dprime, rate = 1)
    } else {0}
```

```
    total <- total + part1*part2
  }
  return(total/denom)
}, vectorize.args = "b")

# Define F_{B_{22}} using numerical integration
F_B22 <- Vectorize(function(upper, n, term1, term2, d, dstar){
  if(upper <= 0){return(0)}
  result <- integrate(
    function(b_val) f_B22(b_val, n, term1, term2, d, dstar),
    lower = 0, upper = upper)$value
  return(result)
}, vectorize.args = 'upper')

L_222 <- function(a, n, term1, term2, d, dstar){
  theta_221 <- (a+(n-d)*term1)/d
  result <- 2*dstar - theta_221*(dstar+d) - (n-dstar)*(term2-term1)
  return(result)
}

U_222 <- function(a, n, term1, term2, d, dstar){
  theta_221 <- (a+(n-d)*term1)/d
  result <- theta_221*(dstar-d) - (n-dstar)*(term2-term1)
  return(result)
```

```
}

# General integrand computation (for pi_2221 or pi_2222)
comp2case22_integrand <- function(a, n, term1, term2,
d, dstar, type = "2221"){
  a1 <- L_222(a, n, term1, term2, d, dstar)
  a2 <- U_222(a, n, term1, term2, d, dstar)

  f1 <- F_B22(a1, n, term1, term2, d, dstar)
  f2 <- F_B22(a2, n, term1, term2, d, dstar)

  fa <- f_A12(a, term1, d)
  if(type == "2221"){
    result <- (f2-f1)*fa
  } else if (type == "2222"){
    result <- (f1-f2)*fa
  }
  return(result)
}

# Used to define pi_2221 or pi_2222
pi_222 <- function(n, term1, term2, d, dstar, type = "2221"){
  if(type == "2221"){
    lower_bd <- d - (n-d)*term1
```

```
    upper_bd <- d*term1
    fit_call <- quote(function(a)
      comp2case22_integrand(a, n, term1, term2,
        d, dstar, type = "2221"))
  } else if (type == "2222"){
    lower_bd <- 0
    upper_bd <- d - (n-d)*term1
    fit_call <- quote(function(a)
      comp2case22_integrand(a, n, term1, term2,
        d, dstar, type = "2222"))
  }
  if(lower_bd < upper_bd){
    result <- integrate(
        eval(fit_call), lower = lower_bd,
        upper = upper_bd)$value
  } else {
    result <- 0
  }
  return(result)
}

# Comparison 2 Case 2 Full Computation
comp2case2 <- function(n, r, term1, term2){
  total <- 0
```

```
  for(d in r:n){
    prob_d <- 1 - exp(-term1)
    pid1 <- dbinom(d, size = n, prob = prob_d)
    total2 <- 0
    for(dstar in d:n){
      dprime <- dstar - d
      prob_dprime <- 1 - exp(-(term2 - term1))
      pid2 <- dbinom(dprime, size = (n-d), prob = prob_dprime)
      if(dstar == d){ # Case 2.1
        value <- pid2 * pi_2212(n, term1, term2, d)
      } else { # Case 2.2
        part1 <- pi_222(n, term1, term2, d, dstar, type = "2221")
        part2 <- pi_222(n, term1, term2, d, dstar, type = "2222")
        value <- pid2 * (part1 + part2)
      }
      total2 <- total2 + value
    }
    total <- total + (pid1 * total2)
  }
  return(total)
}
```

A.3.3 Comparison 2 Case 3

```
# Define f_{B_{231}}(b)
```

```
f_B231 <- function(b, n, r, term1, term2, d){
  tprime <- term2 - term1
  part1 <- (r-d)/((1-exp(-tprime))^(r-d))
  omega <- 1 + (n-r)/2
  total <- 0
  for(i in 0:(r-d-1)){
    taui <- tprime*(omega + i)
    thetai <- (omega + i)/(i + 1)
    part2 <- (-1)^(i) * choose((r-d-1), i)

    part3 <- exp(-(b/thetai))/(thetai*(i+1))
    part4 <- (1 - 1/thetai)^(-(r-d-1))
    incl1 <- (1 - 1/thetai) * b
    part5 <- ifelse(incl1 < 0, 0,
        pgamma(incl1, shape = r - d - 1, lower.tail = TRUE))

    expr1 <- part3*part4*part5

    part6 <- exp(-tprime*(i+1)) / (thetai * (i+1))
    part7 <- exp(-(b-taui)/thetai)
    part8 <- (1 - 1/thetai)^(-(r-d-1))
    incl2 <- (1 - 1/thetai) * (b-taui)
    part9 <- ifelse(incl2 < 0, 0,
        pgamma(incl2, shape = r - d - 1, lower.tail = TRUE))
```

```
    expr2 <- part6*part7*part8*part9

    total <- total + part2*(expr1 - expr2)
  }
  result <- total * part1
  return(result)
}

# Define F_{B_{231}}(b) using numerical integration
F_B231 <- Vectorize(function(upp_bd, n, r, term1, term2, d){
  omega <- 1 + (n-r)/2
  tprime <- term2 - term1
  high_val <- (r-d-1+omega) * tprime
  if(upp_bd >= high_val){
    return(1)
  } else if (upp_bd <= 0){
    return(0)
  } else {
    res <- integrate(
    function(b) f_B231(b, n, r, term1, term2, d),
    lower = 0, upper = upp_bd)$value
  }
  return(res)
```

```
},
vectorize.args = "upp_bd")

U_231 <- function(a, n, r, term1, term2, d){
  part1 <- r - ((n-r)/2)*term2
  part2 <- (r-d-1)*term1
  part3 <- (1 + ((n-r)/2))*term1
  res <- part1 - part2 - part3 - a
  return(res)
}

# Integrand for pi_231
inn_int_231 <- function(a, n, r, term1, term2, d){
  alph <- U_231(a, n, r, term1, term2, d)
  fa <- f_A12(a, term1, d)
  fb <- F_B231(alph, n, r, term1, term2, d)
  return(fa*fb)
}

pi_231 <- function(n, r, term1, term2, d){
  upper_bd <- d*term1
  res <- integrate(
    function(a) inn_int_231(a, n, r, term1, term2, d),
    lower = 0, upper = upper_bd)$value
```

```
  return(res)
}

# Define f_{X_{r:n}}(x)
comp2case3_fxrn <- Vectorize(function(x_r, n, r,
term1, term2, d, dstar){
  if(x_r > term1){
    par1 <- factorial(dstar-d)
    par2 <- (factorial(r-d-1) * factorial(dstar -r))
    part1 <- par1/par2
    denom <- (exp(-term1) - exp(-term2))^(dstar - d)
    num1 <- (exp(-term1) - exp(-x_r))^(r-d-1)
    num2 <- (exp(-x_r) - exp(-term2))^(dstar-r)
    num3 <- exp(-x_r)
    value <- part1 * ((num1 * num2 * num3) / denom)
    return(value)
  } else {
    return(0)
  }
}, vectorize.args = "x_r")

# Define f_{B_{23}}
f_B23 <- Vectorize(function(b, x_r, n, r, term2, dstar){
  if(x_r >= term2) {return(0)}
```

```
  tprime <- term2 - x_r
  d2 <- dstar - r
  if(d2 <= 0){return(0)}
  total <- 0
  denom <- (1 - exp(-tprime))^d2
  for(i in 0:d2){
    part1 <- choose(d2, i) * exp(-i*tprime) * (-1)^i
    q <- b - i * tprime
    part2 <- if(q > 0){
      dgamma(q, shape = d2, rate = 1)
    } else {0}
    total <- total + part1 * part2
  }
  return(total / denom)
}, vectorize.args = "b")

# Define F_{B_{23}} using numerical integration
F_B23 <- Vectorize(function(upper, x_r, n, r, term2, dstar){
  if(upper <= 0){return(0)}
  result <- integrate(
    function(b) f_B23(b, x_r, n, r, term2, dstar),
    lower = 0, upper = upper, subdivisions = 500L,
    stop.on.error = FALSE)$value
  return(result)
```

```
}, vectorize.args = 'upper')

L_232 <- function(a, x_r, n, r, term2, dstar){
  part1 <- (1/r + 1/dstar) * a
  part2 <- ( (n-r+1)/r + (dstar - r +1)/dstar ) * x_r
  part3 <- (n-dstar)*term2 / dstar
  result <- dstar*(2 - part1 - part2 - part3)
  return(result)
}

U_232 <- function(a, x_r, n, r, term2, dstar){
  part1 <- (1/r - 1/dstar) * a
  part2 <- ( (n-r+1)/r - (dstar - r +1)/dstar ) * x_r
  part3 <- (n-dstar)*term2 / dstar
  result <- dstar*(part1 + part2 - part3)
  return(result)
}

# The inner integrand
pi_232_inner_int <- Vectorize(function(a, x_r, n, r, term2,
dstar, type = "2321"){
  a1 <- L_232(a, x_r, n, r, term2, dstar)
  a2 <- U_232(a, x_r, n, r, term2, dstar)
```

```
b_upp_bd <- (term2 - x_r) * (dstar - r)
if(a1 < 0){f1 <- 0}
else if (a1 >= b_upp_bd){f1 <- 1}
else {
  f1 <- F_B23(a1, x_r, n, r, term2, dstar)
  if(f1 > 1){f1 <- 1}
}

if(a2 < 0){f2 <- 0}
else if (a2 >= b_upp_bd){f2 <- 1}
else {
  f2 <- F_B23(a2, x_r, n, r, term2, dstar)
  if(f2 > 1){f2 <- 1}
}

fa <- f_A11(a, x_r, r)
# Safeguarding cases where it's negative but close to 0
fa[fa < 0 & abs(fa) < 1e-10] <- 0

if(type == "2321"){subtra = (f2-f1)}
else if (type == "2322"){subtra = (f1-f2)}
# Safeguarding once again
subtra[subtra < 0 & abs(subtra) < 3e-4] <- 0
result = subtra*fa
```

```
  return(result)
}, vectorize.args = "a")

# The outer integrand for pi_2321 and pi_2322
pi_232_outer_int <- Vectorize(function(x_r, n, r, term1,
term2, d, dstar, type = "2321"){
  if(type == "2321"){
    upper_bd <- (r-1)*x_r
    lower_bd <- r - x_r*(n-r+1)
    if(lower_bd <= 0){lower_bd <- 0}
    fit_call <- quote(function(a)
      pi_232_inner_int(a, x_r, n, r, term2, dstar, type = "2321"))
  } else if (type == "2322"){
    upper_bd <- r - x_r*(n-r+1)
    lower_bd <- 0
    fit_call <- quote(function(a)
      pi_232_inner_int(a, x_r, n, r, term2, dstar, type = "2322"))
  }
  if(lower_bd < upper_bd){
    result <- integrate(
        eval(fit_call), lower = lower_bd, upper = upper_bd,
        subdivisions = 500L, stop.on.error = FALSE)$value
  } else {result <- 0}
```

```
  fxr <- comp2case3_fxrn(x_r, n, r, term1, term2, d, dstar)
  return(result*fxr)
}, vectorize.args = "x_r")

# Comparison 2 Case 3 Full Computation
comp2case3 <- function(n, r, term1, term2){
  total <- 0
  prob_d <- 1 - exp(-term1)
  prob_dprime <- 1 - exp(-(term2 - term1))
  for(d in 0:(r-1)){
    pid1 <- dbinom(d, size = n, prob = prob_d)
    total2 <- 0
    for(dstar in r:n){
      dprime <- dstar - d
      pid2 <- dbinom(dprime, size = (n-d), prob = prob_dprime)
      if(dstar == r){
        pi_232_r <- pi_231(n, r, term1, term2, d)
        total2 <- total2 + (pid2 * pi_232_r)
      } else {
        pi_2321 <- integrate(
          function(x_r) pi_232_outer_int(x_r, n, r, term1,
            term2, d, dstar, type = "2321"),
          lower = term1, upper = term2)$value
```

```
        pi_2322 <- integrate(
          function(x_r) pi_232_outer_int(x_r, n, r, term1,
            term2, d, dstar, type = "2322"),
          lower = term1, upper = term2)$value
        total2 <- total2 + (pid2 * (pi_2321 + pi_2322))
      }
    }
    total <- total + (pid1 * total2)
  }
  return(total)
}
```

A.4 Demonstration of Code Usage

```
# Comparison 1 Example
c1 <- comp1case1(n = 15, r = 4, s = 6, term1 = 0.75)
c2 <- comp1case2(n = 15, r = 4, s = 6, term1 = 0.75)
c3 <- comp1case3(n = 15, s = 6, term1 = 0.75)
c1 + c2 + c3

# Comparison 2 Example
c1 <- comp2case1(n = 10, r = 4, term1 = 0.5, term2 = 0.75)
c2 <- comp2case2(n = 10, r = 4, term1 = 0.5, term2 = 0.75)
c3 <- comp2case3(n = 10, r = 4, term1 = 0.5, term2 = 0.75)
c1 + c2 + c3
```

Bibliography

[1] H. A. Almuzaini and M. Z. Raqab. Comparative study for assessing the pitman's closeness of predictors based on exponential record data. *Communications in Statistics-Simulation and Computation*, 51(8):4427–4449, 2022. doi:10.1080/03610918.2020.1743857.

[2] A. E. Aly. Prediction of the exponential fractional upper record-values. *Mathematica Slovaca*, 72(2):491–506, 2022. doi:10.1515/ms-2022-0032.

[3] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A first course in order statistics*. SIAM, 2008.

[4] N. Balakrishnan and R. Aggarwala. *Progressive censoring: theory, methods, and applications*. Springer Science & Business Media, 2000.

[5] N. Balakrishnan and A. C. Cohen. *Order statistics & inference: estimation methods*. Elsevier, 2014.

[6] N. Balakrishnan and E. Cramer. *Art of progressive censoring*. Springer, 2014.

[7] N. Balakrishnan and K. F. Davies. Pitman closeness results for type-I censored data from exponential distribution. *Statistics & Probability Letters*, 83(12):2693–2698, 2013. doi:10.1016/j.spl.2013.09.007.

[8] N. Balakrishnan, K. F. Davies, J. P. Keating, and R. L. Mason. Pitman closeness, monotonicity and consistency of best linear unbiased and invariant estimators for exponential distribution under type-II censoring. *Journal of Statistical Computation and Simulation*, 81(8):985–999, 2010. doi:10.1080/00949651003587684.

[9] N. Balakrishnan, E. Cramer, and D. Kundu. *Hybrid censoring know-how: Designs and implementations*. Academic Press, 2023.

[10] S.-M. Chen and G. K. Bhattacharyya. Exact confidence bounds for an exponential parameter under hybrid censoring. *Communications in Statistics-Theory and Methods*, 16(8):2429–2442, 1987. doi:10.1080/03610928708829516.

[11] A. Childs, B. Chandrasekar, N. Balakrishnan, and D. Kundu. Exact likelihood inference based on type-I and type-II hybrid censored samples from the exponential distribution. *Annals of the Institute of Statistical Mathematics*, 55:319–330, 2003. doi:10.1007/BF02530502.

[12] K.-C. Chiou, T.-H. Huang, K.-S. Chen, and C.-M. Yu. Fuzzy evaluation model for lifetime performance using type-I censoring data. *Mathematics*, 12(13), 2024. ISSN 2227-7390. doi:10.3390/math12131935.

[13] K. F. Davies. Pitman closeness results for type-I hybrid censored data from exponential distribution. *Journal of Statistical Computation and Simulation*, 91 (1):58–80, 2020. doi:10.1080/00949655.2020.1806280.

[14] R. Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[15] B. Epstein. Truncated life tests in the exponential case. *The Annals of Mathematical Statistics*, pages 555–564, 1954. doi:10.1214/aoms/1177728723.

[16] S. Karlin. Admissibility for estimation with quadratic loss. *The Annals of Mathematical Statistics*, 29(2):406–436, 1958.

[17] J. P. Keating and R. L. Mason. Practical relevance of an alternative criterion in estimation. *The American Statistician*, 39(3):203–205, 1985.

[18] J. P. Keating, R. L. Mason, and P. K. Sen. *Pitman's measure of closeness: a comparison of statistical estimators*. SIAM, 1993.

[19] J. F. Lawless. *Statistical models and methods for lifetime data*. John Wiley & Sons, 2003.

[20] K. Pearson. Method of moments and method of maximum likelihood. *Biometrika*, 28(1/2):34–59, 1936. doi:10.2307/2334123.

[21] E. J. Pitman. The "closest" estimates of statistical parameters. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 33. Cambridge University Press, 1937. doi:10.1017/S0305004100019563.

[22] C. R. Rao. Some comments on the minimum mean square error as a criterion of estimation. Technical report, 1980.

[23] R. J. Serfling. *Approximation theorems of mathematical statistics*. John Wiley & Sons, 2009.

[24] M.-J. Sun, Y.-C. Zhang, F.-R. Lin, S. Wang, L.-W. Liu, and J.-L. Qu. Rapid fluorescence lifetime imaging microscopy via few-photon imaging. *APL Photonics*, 9(1), 2024. doi:10.1063/5.0178452.