University of Tehran,
College of Engineering

Homework 2

Introduction to Statistical Interference

Instructor: **Dr. Mohammadreza A. Dehaqani**

Erfan Panahi (*Student Number:* **810103084**)

Problem 1. Oil Pipeline Pressure Monitoring

Pressure measurements, $X_1, X_2, \ldots, X_n$, satisfy the following model:

$$X_i = \mu + \epsilon_i$$

where μ is the unknown true average pressure, and ϵ_i represents random error. The errors are i.i.d. with mean 0 and unknown standard deviation σ.

The pipeline's pressure is measured 100 times. The recorded mean pressure is 75,348 Pascals, with a standard deviation of 25 Pascals.

$$\sigma_{\bar{X}_{10}} = \frac{\sigma}{\sqrt{n}} = \frac{25}{10} = 2.5\,, \qquad \bar{X}_{10} = 75{,}348$$

Part a. Now we drive an approximate 95% confidence interval for μ.

$$P\left(\bar{X}_{10} - z_{\frac{\alpha}{2}}\sigma_{\bar{X}_{10}} \leq \mu \leq \bar{X}_{10} + z_{\frac{\alpha}{2}}\sigma_{\bar{X}_{10}}\right) = 1 - \alpha$$

$$\alpha = 0.05\,, \qquad z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$$

$$\rightarrow 75{,}348 - 1.96 \times 2.5 \leq \mu \leq 75{,}348 + 1.96 \times 2.5$$

$$\mathbf{75343.1 \leq \boldsymbol{\mu} \leq 75352.9}$$

Part b. Now we check each statement:

1. **To estimate the average of the 100 pressure measurements and give ourselves some room for error in the estimate.**

   This is **false** because the interval is constructed to estimate the population parameter (μ), not the sample statistic.

2. **To estimate the true average pressure of the pipeline and give ourselves some room for error in the estimate.**

   This is **correct**. The confidence interval provides a range of values within which the true average pressure of the pipeline (μ) is likely to fall.

3. **To provide a range in which 95 of the 100 pressure measurements are likely to have fallen.**

   This is **false** because the confidence interval is for the mean of the population, not for individual measurements.

4. **To provide a range in which 95% of all possible pressure measurements are likely to fall.**

   This is **false** as well, for the same reason as above. The interval does not cover individual measurements but estimates the population mean μ.

Part c. With the given information, it's not possible to sketch a precise histogram of the 100 measurements. We only have the sample mean and standard deviation, which gives us some idea about the central tendency and spread of the data, but not the **actual distribution** of the individual measurements.

To create a histogram, we would need the raw data points or at least some information about the distribution of the errors (ϵ_i). Without knowing the error distribution, we can't accurately represent the shape or spread of the data in a histogram.

Part d. To determine the number of pressure measurements (n) needed to ensure that the average pressure is within 1 Pascal of the true pressure with 95% confidence, we need to use the formula for the margin of error in the context of a confidence interval:

$$\text{Margin of error} = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\rightarrow \text{Margin of error} = 1 \rightarrow z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \times \frac{25}{\sqrt{n}} \rightarrow n = 25^2 \times 1.96^2 = 2401$$

So, the engineer would need to record approximately $\mathbf{2,461}$ pressure measurements to ensure that the average pressure is within 1 Pascal of the true pressure with 95% confidence.

Problem 2. Manufacturing Quality Control

Part a. First, we identify the given parameters of the problem:

- $n = 625$
- Confidence interval for the mean strength (Newtons): $(126.45, 128.55)$

$$\text{Margin of error} = \frac{128.55 - 126.45}{2} = 1.05$$

The confidence interval formula for the mean strength is:

$$\text{CI for } \mu = \bar{x} \pm z\frac{\sigma}{\sqrt{n}}$$

$$\text{Margin of error} = z\frac{\sigma}{\sqrt{n}}$$

Now we estimate the population mean (μ) and standard deviation (σ).

$$\mu = \frac{128.55 + 126.45}{2} = 127.5$$

For 99% confidence, $z \approx 2.574$. Rearrange to solve for σ:

$$\sigma = \frac{1.05\sqrt{n}}{z} = \frac{1.05\sqrt{625}}{2.576} = \frac{26.25}{2.576} \approx 10.1942$$

The proportion of defective springs is determined by the area under the normal curve beyond 140 (N):

$$z = \frac{140 - \mu}{\sigma} = \frac{140 - 127.5}{10.1942} \approx 1.2262$$

From the standard normal table, the area to the right of $z = 1.23$ (proportion defective):

$$\hat{p} = P(Z > 1.2262) \approx 0.1101$$

Use the normal approximation for proportions to construct the confidence interval:

$$\sigma_n = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.1101(1-0.1101)}{625}} \approx 0.0125$$

For 90% confidence, $z = 1.645$:

$$\text{Margin of Error} = ME = z \times \sigma_n = 1.645 \times 0.0125 \approx 0.0206$$

Confidence interval:

$$(\hat{p} - ME, \hat{p} + ME) = (0.1101 - 0.0206, 0.1101 + 0.0206) \approx (0.0895, 0.1307) = (\mathbf{8.95}\%, \mathbf{13.07}\%)$$

Part b. Can the confidence interval for the percentage of defective springs be computed?

Yes, the confidence interval for the percentage of defective springs can be computed because the problem provides sufficient information to estimate:

1. The population mean strength (μ) from the midpoint of the given confidence interval.
2. The population standard deviation (σ) indirectly using the margin of error and sample size.

3. The proportion of defective springs using the cumulative probability of the normal distribution above $140\ (N)$.

Therefore, we can construct the confidence interval as shown in part (a).

Problem 3. Ancient War between Persians and Greeks

The Law of Large Numbers (LLN) holds if, as $n \to \infty$:

$$P\left(\left|\frac{Sn}{n} - E\left[\frac{S_n}{n}\right]\right| < \epsilon\right) \to 0$$

where S_n is the total number of soldiers reaching the fortresses out of sent, and $E\left[\frac{S_n}{n}\right] = p$, the expected proportion of success per soldier.

The key is whether $\frac{S_n}{n}$ (the average success rate) converges in probability to its expected value p as $n \to \infty$. This depends on whether the success events are independent and identically distributed (i.i.d.).

case a. Each soldier is sent through a completely different route to the fortress.

Each soldier's success depends on whether their unique route is blocked. The blocking of routes is independent, so each soldier has an independent success probability p. Since the soldiers' success probabilities are i.i.d., the average success rate $\frac{S_n}{n}$ converges to p by the LLN. (LLN holds)

case b. The soldiers are split into $\frac{n}{2}$ pairs, and each pair is sent through its own route.

For each pair, the two soldiers' success depends on the same route being unblocked. If the route is unblocked, both succeed; if blocked, both fail. The success of pairs is independent, as different pairs use different routes, and the probability of success is p. Within each pair, the soldiers' outcomes are perfectly correlated, but this does not affect $\frac{S_n}{n}$ because success remains i.i.d. at the pair level. As $n \to \infty$, $\frac{S_n}{n}$ still converges to p. (LLN holds)

case c. The soldiers are split into two groups of $\frac{n}{2}$, and each group is sent through the same route, with the two groups sent through different routes.

Within each group, all $\frac{n}{2}$ soldiers succeed or fail together, depending on whether their route is blocked. The two groups' outcomes are independent, but within each group, soldiers' outcomes are perfectly correlated. The variance of $\frac{S_n}{n}$ does not decrease as $n \to \infty$, because the outcomes are not independent at the individual soldier level. Since the LLN requires individual outcomes to be i.i.d., and here they are not, LLN does not hold.

case d. All the soldiers are sent through one route.

If the route is unblocked, all n soldiers succeed; if blocked, all soldiers fail. The outcome for all n soldiers is perfectly correlated: either all *succeed* or all *fail*. The variance of $\frac{S_n}{n}$ does not decrease as $n \to \infty$, and the proportion of success does not converge to p. (LLN does not hold)

Problem 4. Estimating π by Throwing Darts

Part a. To estimate π using dart throws, we can apply statistical principles from the law of large numbers and confidence intervals. Here's the outline of the solution:

Each dart throw is a Bernoulli trial with success probability $\frac{\pi}{4}$ (the area ratio of the circle to the square). The random variable X_i for each dart throw takes values 1 (success, landing inside the circle) or 0 (failure, landing outside). The proportion of darts that land inside the circle, $\hat{p} = \frac{number\ of\ sucesses}{total\ trials}$, is an estimator for $p = \frac{\pi}{4}$. Then $\hat{\pi} = 4\hat{p}$, is an estimator for π.

Each dart throw is a Bernoulli trial, where $X_i = 1$ with probability $p = \frac{\pi}{4}$, and $X_i = 0$ with probability $1 - p = 1 - \frac{\pi}{4}$. The sample mean $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ provides an estimate of $\frac{\pi}{4}$. From $\bar{X}_n$, we estimate π as $\hat{\pi} = 4\bar{X}_n$. The estimation error is $|\hat{\pi} - \pi|$, which we rewrite in terms of $\bar{X}_n$:

$$|\hat{\pi} - \pi| = |4\bar{X}_n - \pi| = 4\left|\bar{X}_n - \frac{\pi}{4}\right|$$

We need the estimation error to be no more than 0.01 with at least 95% probability. Therefore, we solve for n such that:

$$P\left(4\left|\bar{X}_n - \frac{\pi}{4}\right| \le 0.01\right) \ge 0.95 \rightarrow P\left(\left|\bar{X}_n - \frac{\pi}{4}\right| \le 0.0025\right) \ge 0.95$$

Using the normal approximation:

$$P\left(-0.0025 \le \bar{X}_n - \frac{\pi}{4} \le 0.0025\right) \approx P\left(-0.0025 \le Z\sqrt{\frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n}} \le 0.0025\right)$$

$$= P\left(-\frac{0.0025}{\sqrt{\frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n}}} \le Z \le \frac{0.0025}{\sqrt{\frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n}}}\right)$$

Here, Z is the standard normal random variable. The cumulative probability of Z between $-z$ and z for 95% confidence is:

$$z = 1.96 \rightarrow 1.96 \ge \frac{0.0025}{\sqrt{\frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n}}} \rightarrow n \ge \frac{1.96^2 \times \frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{0.0025^2} \approx 103598.9 \rightarrow \boldsymbol{n = 103599}$$

Problem 5. Parameter Estimation for an Exponential Distribution

The given probability density function (PDF) is

$$f(x;\alpha) = \frac{x}{\alpha^2} e^{-\frac{x}{\alpha}}, \qquad x > 0\,, \alpha > 0$$

Part a. For a random sample $X_1, X_2, \dots, X_n$, the likelihood function is:

$$L(\alpha) = \prod_{i=1}^{n} f(x_i; n) = \prod_{i=1}^{n} \frac{x_i}{\alpha^2} e^{-\frac{x_i}{\alpha}}$$

Take the log of the likelihood function to obtain the log-likelihood:

$$\ell(\alpha) = \mathrm{Ln}\big(L(\alpha)\big) = \sum_{i=1}^{n} \mathrm{Ln}\left(\frac{x_i}{\alpha^2} e^{-\frac{x_i}{\alpha}}\right) = \sum_{i=1}^{n} \left(\mathrm{Ln}(x_i) - 2\mathrm{Ln}(\alpha) - \frac{x_i}{\alpha}\right)$$

To maximize $\ell(\alpha)$, take the derivative with respect to α and set it to zero:

$$\frac{\partial \ell(\alpha)}{\partial \alpha} = \sum_{i=1}^{n} \left(\frac{x_i}{\alpha^2} - \frac{2}{\alpha}\right) = \frac{\sum_{i=1}^{n} x_i}{\alpha^2} - \frac{2n}{\alpha} = 0 \rightarrow \alpha = \frac{\sum_{i=1}^{n} x_i}{2n}$$

where $\frac{\sum_{i=1}^{n} x_i}{n}$ is the sample mean ($\bar{x}$).

Thus, the MLE for α is:

$$\hat{\alpha} = \frac{\bar{x}}{2}$$

Now we calculate the mean sample and the parameter α:

$$x_1 = 0.25\,, \qquad x_2 = 0.75\,, \qquad x_3 = 1.50\,, \qquad x_4 = 2.50\,, \qquad x_5 = 2.00$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{0.25 + 0.75 + 1.5 + 2.5 + 2}{5} = \frac{7}{5} = 1.4$$

$$\hat{\alpha} = \frac{\bar{x}}{2} = \frac{1.4}{2} = 0.7$$

The maximum likelihood estimate (MLE) for α is:

$$\boldsymbol{\hat{\alpha}_{MLE} = 0.7}$$

Part b. The expected value (mean) of X is:

$$E[X] = \int_{-\infty}^{+\infty} x f(x;\alpha) dx = \int_{0}^{+\infty} \frac{x^2}{\alpha^2} e^{-\frac{x}{\alpha}} dx = 2\alpha$$

The first population moment is $E[X] = 2\alpha$. The first sample moment is the sample mean $\bar{x}$. Using the method of moments:

$$E[X] = 2\alpha = \bar{x} \rightarrow \hat{\alpha}_{MoM} = \frac{\bar{x}}{2} = 0.7$$

The method of moments (MoM) estimate for α is:

$$\boldsymbol{\hat{\alpha}_{MoM} = 0.7}$$

Problem 6. Roulette Simulation and Profit Analysis

Part 1, 2. The simulation for N-rounds involves betting 1 dollar on black. If the outcome is black, the bet is won, and the earnings increase by 1 dollar. If the outcome is red or green, the bet is lost, and the earnings decrease by 1 dollar.

To study the distribution of total earnings for different values of N, we performed Monte Carlo simulations with 100,000 rounds for each $\boldsymbol{N}$ $(10, 25, 100, 1000)$. The results were plotted in figure 6.1 to analyze the distributions.

As N increases, the distributions resemble a normal distribution due to the **Central Limit Theorem (CLT)**, with a mean of $E[S_N] = N.E[X]$ and standard deviation $var(S_N) = N \cdot Var(X)$.

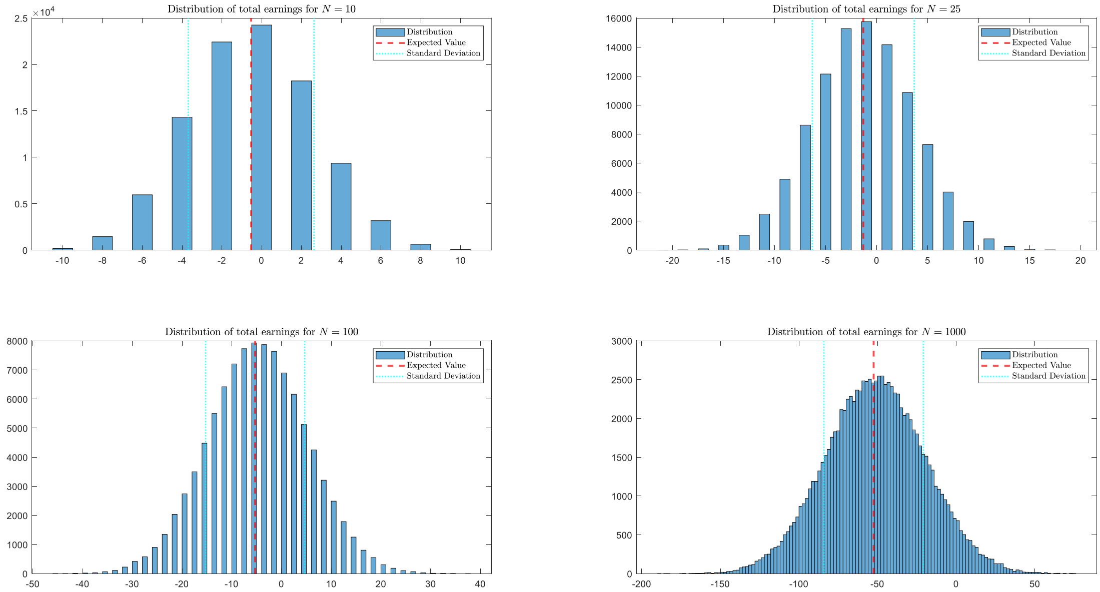


Figure 6.1. The theorical results for different number of N

Part 3. Similar to the previous analysis, we performed Monte Carlo simulations to study the distribution of average winnings for different values of N. The results were plotted in figure 3.2 to observe changes in expected values and standard errors.

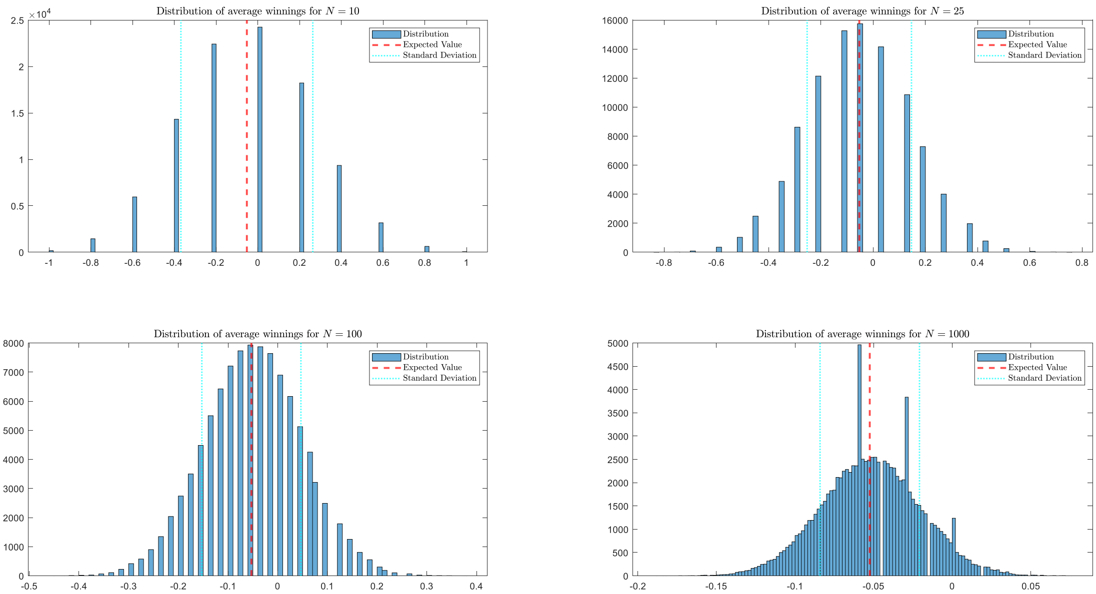


Figure 6.2. The theorical results for different number of N

Part 4.

- **Simulation Results:** The simulation results are shown in figure 3.3.

```
N = 10:      E[Sn] = -0.531280,     sigma_N = 3.170464
N = 25:      E[Sn] = -1.301500,     sigma_N = 5.006710
N = 100:     E[Sn] = -5.281440,     sigma_N = 9.981140
N = 1000:    E[Sn] = -52.666660,    sigma_N = 31.437806
```

Figure 6.3. The theorical results for different number of N

- **Theorical:** Now we should calculate the expected values and standard errors of S_N for each N. We'll compare these theoretical values with the Monte Carlo simulation results.

  Expected Value ($E(S_N)$): The expected value of S_N after N rounds, where you bet 1 dollar on black each round, is given by:

$$E(S_N) = N \cdot E(X) = N \cdot \big(p \times 1 + (1-p) \times (-1)\big) = (2p-1) \cdot N = -\frac{N}{19}$$

  where $E(X)$ is the expected value for one round.
  Variance for N rounds (σ_N^2): The variance for one round is given by:

$$\sigma_N^2 = N \cdot \sigma = N \cdot (p \times 1^2 + (1-p) \times (-1)^2 - (2p-1)^2) = N \cdot (1-p^2) = (1-p^2)N = \frac{360}{361}N$$

  Thus, the standard error (σ_N) will be:

$$\sigma_N = \sqrt{\frac{360}{361}N} = \frac{6\sqrt{10N}}{19}$$

  The comparison of theorical and simulation results for different number of N are shown in Table 6.1.

N	$E(S_N)$ simulation	$E(S_N)$ theorical	σ_N^2 simulation	σ_N^2 theorical
10	-0.531280	-0.526	3.170464	3.162
25	-1.301500	-1.315	5.006710	5
100	-5.281440	-5.26	9.981140	10
1000	-52.666660	-52.6	31.43806	31.62

Table 6.1. The comparison of theorical and simulation results for different number of N

Part 5.

- **Simulation Results:** Using the previous simulation results we calculate the probability that the sum is less than or equal to 0. This probability is shown in figure 6.4.

```
losing_rate = 0.6046
```

Figure 6.4. The theorical results for different number of N

- **Theorical:** Using the standard normal distribution, we calculate the probability that the sum is less than or equal to 0.

$$P(S_{25} \le 0) = P\left(\frac{S_{25} - E(S_{25})}{\sigma_{25}} \le \frac{0-(-1.315)}{5}\right) = P\left(Z \le \frac{1.315}{5}\right) \approx P(Z \le 0.263)$$

  Using the standard normal table or a calculator, we find:

$$P(Z \le 0.263) \approx 0.603$$

  As we see the theorical and simulation results are approximately the same.

Part 6. The probability of the casino losing money was plotted in figure 6.5 as a function of N for values ranging from 25 to 1000. This analysis helps understand why casinos might encourage players to continue betting.

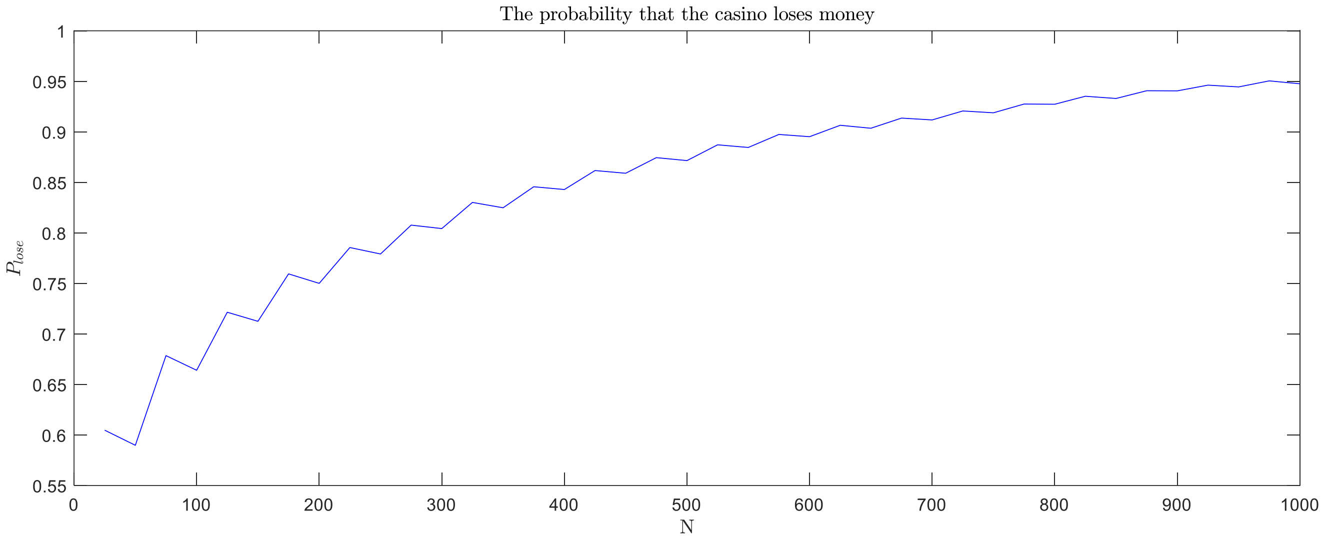


Figure 6.5. The probability of the casino losing money as a function of N for values ranging from 25 to 1000

Problem 8. Convergence in a Noisy Measurement Process

Each estimator X_n is defined as:

$$X_n = \theta + \frac{Z_n}{\sqrt{n}}, \qquad Z_n \sim \mathcal{N}(0,1)$$

Part a. Mean-Square Convergence

To determine mean-square convergence, we need to calculate $E[(X_n - \theta)^2]$.

$$E[(X_n-\theta)^2] = E\left[\left(\theta + \frac{Z_n}{\sqrt{n}} - \theta\right)^2\right] = E\left[\left(\frac{Z_n}{\sqrt{n}}\right)^2\right] = \frac{1}{n}E[Z_n^2]$$

Since $Z_n \sim \mathcal{N}(0,1)$, we know that:

$$E[Z_n^2] = var(Z_n) + (E[Z_n])^2 = 1 + 0 = 1 \to E[(X_n-\theta)^2] = \frac{1}{n}$$

To determine **mean-square convergence**, we need to check:

$$\lim_{n\to\infty} E[(X_n-\theta)^2] = \lim_{n\to\infty}\frac{1}{n} = 0$$

$$\boldsymbol{X_n \xrightarrow{m.s} \theta}$$

Part b. Convergence in Probability

To determine convergence in probability, we need to analyze:

$$\lim_{n\to\infty} P(|X_n-\theta| > \epsilon) = 0$$

First, we express $|X_n - \theta|$:

$$P(|X_n-\theta|>\epsilon) = P\left(\left|\frac{Z_n}{\sqrt{n}}\right| > \epsilon\right) = P\big(|Z_n| > \epsilon\sqrt{n}\big) = 2P\big(Z_n > \epsilon\sqrt{n}\big) \approx \frac{1}{\epsilon\sqrt{2\pi n}}\exp\left(-\frac{\epsilon^2 n}{2}\right)$$

As n approaches infinity, the exponential term $\exp\left(-\frac{\epsilon^2 n}{2}\right)$ and $\frac{1}{\sqrt{n}}$ dominates and goes to zero very rapidly. Thus:

$$\lim_{n\to\infty} P(|X_n-\theta|>\epsilon) = \lim_{n\to\infty}\frac{1}{\epsilon\sqrt{2\pi n}}\exp\left(-\frac{\epsilon^2 n}{2}\right) = 0$$

Therefore, X_n converges to θ in probability.

$$\boldsymbol{X_n \xrightarrow{p} \theta}$$

Part c. Convergence in Distribution

To determine if X_n converges to θ in distribution, we need to analyze the distribution of $X_n - \theta$ as $n \to \infty$. We have:

$$X_n - \theta = \frac{Z_n}{\sqrt{n}}$$

As $n \to \infty$, $\frac{Z_n}{\sqrt{n}} \to 0$ in distribution because the distribution of $\frac{Z_n}{\sqrt{n}}$ becomes more and more concentrated around 0.

Therefore, X_n converges to θ\theta in distribution, as the distribution of $X_n - \theta$ approaches the distribution of a constant (which is zero).

$$\boldsymbol{X_n \xrightarrow{d} \theta}$$

Problem 9. Failure Rates in a Model with Time-Dependent Scaling

Part a. In this part we should derive the log-likelihood function $\ell(\alpha, \beta_0, \beta_1)$ for the modified Gamma-Weibull distribution, given the PDF:

$$f_T\big(t; \alpha, \beta(t)\big) = \frac{1}{\Gamma(\alpha)}\left(\frac{t}{\beta(t)}\right)^{\alpha-1} \exp\left(-\frac{t}{\beta(t)}\right), \qquad \beta(t) = \beta_0 + \beta_1 t$$

Given n independent observations $T_1, T_2, \dots, T_n$, the joint PDF is:

$$f(t_1, t_2, \dots, t_n; \alpha, \beta_0, \beta_1) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)}\left(\frac{t_i}{\beta(t_i)}\right)^{\alpha-1} \exp\left(-\frac{t_i}{\beta(t_i)}\right)$$

Taking the natural logarithm of the joint PDF to get the log-likelihood function:

$$\ell(\alpha, \beta_0, \beta_1) = \ln\left(\prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)}\left(\frac{t_i}{\beta(t_i)}\right)^{\alpha-1} \exp\left(-\frac{t_i}{\beta(t_i)}\right)\right) = \sum_{i=1}^{n} \ln\left(\frac{1}{\Gamma(\alpha)}\left(\frac{t_i}{\beta(t_i)}\right)^{\alpha-1} \exp\left(-\frac{t_i}{\beta(t_i)}\right)\right)$$

$$= \sum_{i=1}^{n}\left[\ln\left(\frac{1}{\Gamma(\alpha)}\right) + (\alpha-1)\ln\left(\frac{t_i}{\beta(t_i)}\right) + \ln\left(\exp\left(-\frac{t_i}{\beta(t_i)}\right)\right)\right]$$

$$= n\ln\left(\frac{1}{\Gamma(\alpha)}\right) + (\alpha-1)\sum_{i=1}^{n}\ln\left(\frac{t_i}{\beta(t_i)}\right) - \sum_{i=1}^{n}\frac{t_i}{\beta(t_i)}$$

$$\beta(t) = \beta_0 + \beta_1 t \rightarrow \ell(\alpha, \beta_0, \beta_1) = -n\ln\big(\Gamma(\alpha)\big) + (\alpha-1)\sum_{i=1}^{n}\ln\left(\frac{t_i}{\beta_0 + \beta_1 t_i}\right) - \sum_{i=1}^{n}\frac{t_i}{\beta_0 + \beta_1 t_i}$$

$$= -n\ln\big(\Gamma(\alpha)\big) + (\alpha-1)\sum_{i=1}^{n}[\ln(t_i) - \ln(\beta_0 + \beta_1 t_i)] - \sum_{i=1}^{n}\frac{t_i}{\beta_0 + \beta_1 t_i}$$

$$\boldsymbol{\ell(\alpha, \beta_0, \beta_1) = -n\ln\big(\Gamma(\alpha)\big) + (\alpha-1)\sum_{i=1}^{n}[\ln(t_i) - \ln(\beta_0 + \beta_1 t_i)] - \sum_{i=1}^{n}\frac{t_i}{\beta_0 + \beta_1 t_i}}$$

Part b. Now, we compute the first-order conditions (score functions) by taking the partial derivatives of the log-likelihood function with respect to α, β_0, and β_1. These conditions will help us set up the system of equations for the Maximum Likelihood Estimates (MLE) α, β_0, and β_1.

- Partial Derivative with respect to α:

$$\frac{\partial\ell(\alpha, \beta_0, \beta_1)}{\partial\alpha} = -n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}[\ln(t_i) - \ln(\beta_0 + \beta_1 t_i)] = 0$$

- Partial Derivative with respect to β_0:

$$\frac{\partial\ell(\alpha, \beta_0, \beta_1)}{\partial\beta_0} = (\alpha-1)\sum_{i=1}^{n}\left(-\frac{1}{\beta_0 + \beta_1 t_i}\right) + \sum_{i=1}^{n}\frac{t_i}{(\beta_0 + \beta_1 t_i)^2} = 0$$

- Partial Derivative with respect to β_1:

$$\frac{\partial\ell(\alpha, \beta_0, \beta_1)}{\partial\beta_0} = (\alpha-1)\sum_{i=1}^{n}\left(-\frac{t_i}{\beta_0 + \beta_1 t_i}\right) + \sum_{i=1}^{n}\frac{t_i^2}{(\beta_0 + \beta_1 t_i)^2} = 0$$

Thus, the first-order conditions give us the following system of equations to solve for the MLEs α, β_0, and β_1:

1. $-n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \sum_{i=1}^{n}[\ln(t_i) - \ln(\beta_0+\beta_1 t_i)] = 0$
2. $(\alpha-1)\sum_{i=1}^{n}\left(-\frac{1}{\beta_0+\beta_1 t_i}\right) + \sum_{i=1}^{n}\frac{t_i}{(\beta_0+\beta_1 t_i)^2} = 0$
3. $(\alpha-1)\sum_{i=1}^{n}\left(-\frac{t_i}{\beta_0+\beta_1 t_i}\right) + \sum_{i=1}^{n}\frac{t_i^2}{(\beta_0+\beta_1 t_i)^2} = 0$

Part c. The Fisher Information matrix $I(\alpha,\beta_0,\beta_1)$ consists of the expected values of these second-order partial derivatives:

$$\boldsymbol{I(\alpha,\beta_0,\beta_1)} = \begin{bmatrix} \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\alpha^2}\right] & \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\alpha\partial\beta_0}\right] & \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\alpha\partial\beta_1}\right] \\ \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\alpha\partial\beta_0}\right] & \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\beta_0^2}\right] & \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\beta_0\partial\beta_1}\right] \\ \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\alpha\partial\beta_1}\right] & \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\beta_0\partial\beta_1}\right] & \boldsymbol{E}\left[-\frac{\partial^2\ell}{\partial\beta_1^2}\right] \end{bmatrix}$$

$$\boldsymbol{\frac{\partial^2\ell}{\partial\alpha^2}} = \frac{\partial}{\partial\alpha}\left(-n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right) = -n\left(\frac{\Gamma''(\alpha)\Gamma(\alpha) - \left(\Gamma'(\alpha)\right)^2}{\Gamma(\alpha)^2}\right)$$

$$\boldsymbol{\frac{\partial^2\ell}{\partial\alpha\partial\beta_0}} = \frac{\partial}{\partial\beta_0}\left(-\sum_{i=1}^{n}\ln(\beta_0+\beta_1 t_i)\right) = -\sum_{i=1}^{n}\frac{1}{\beta_0+\beta_1 t_i}$$

$$\boldsymbol{\frac{\partial^2\ell}{\partial\alpha\partial\beta_1}} = \frac{\partial}{\partial\beta_1}\left(-\sum_{i=1}^{n}\ln(\beta_0+\beta_1 t_i)\right) = -\sum_{i=1}^{n}\frac{t_i}{\beta_0+\beta_1 t_i}$$

$$\boldsymbol{\frac{\partial^2\ell}{\partial\beta_0^2}} = \frac{\partial}{\partial\beta_0}\left((\alpha-1)\sum_{i=1}^{n}\left(-\frac{1}{\beta_0+\beta_1 t_i}\right) + \sum_{i=1}^{n}\frac{t_i}{(\beta_0+\beta_1 t_i)^2}\right)$$
$$= (\alpha-1)\sum_{i=1}^{n}\left(\frac{1}{(\beta_0+\beta_1 t_i)^2}\right) - 2\sum_{i=1}^{n}\left(\frac{t_i}{(\beta_0+\beta_1 t_i)^3}\right)$$

$$\boldsymbol{\frac{\partial^2\ell}{\partial\beta_0\partial\beta_1}} = \frac{\partial}{\partial\beta_1}\left((\alpha-1)\sum_{i=1}^{n}\left(-\frac{1}{\beta_0+\beta_1 t_i}\right) + \sum_{i=1}^{n}\frac{t_i}{(\beta_0+\beta_1 t_i)^2}\right)$$
$$= (\alpha-1)\sum_{i=1}^{n}\left(\frac{t_i}{(\beta_0+\beta_1 t_i)^2}\right) - 2\sum_{i=1}^{n}\left(\frac{t_i^2}{(\beta_0+\beta_1 t_i)^3}\right)$$

$$\boldsymbol{\frac{\partial^2\ell}{\partial\beta_1^2}} = \frac{\partial}{\partial\beta_1}\left((\alpha-1)\sum_{i=1}^{n}\left(-\frac{t_i}{\beta_0+\beta_1 t_i}\right) + \sum_{i=1}^{n}\frac{t_i^2}{(\beta_0+\beta_1 t_i)^2}\right)$$
$$= (\alpha-1)\sum_{i=1}^{n}\left(\frac{t_i^2}{(\beta_0+\beta_1 t_i)^2}\right) - 2\sum_{i=1}^{n}\left(\frac{t_i^3}{(\beta_0+\beta_1 t_i)^3}\right)$$

Part d. Suppose that the Fisher Information matrix $I(\alpha,\beta_0,\beta_1)$ is:

$$I(\alpha, \beta_0, \beta_1) = \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\beta_0} & I_{\alpha\beta_1} \\ I_{\beta_0\alpha} & I_{\beta_0\beta_0} & I_{\beta_0\beta_1} \\ I_{\beta_1\alpha} & I_{\beta_1\beta_0} & I_{\beta_1\beta_1} \end{bmatrix}$$

To find the **inverse $\boldsymbol{I^{-1}(\alpha, \beta_0, \beta_1)}$**, we use the formula for the inverse of a 3×3 matrix:

$$I^{-1} = \frac{1}{\det(I)} . \operatorname{adj}(I)$$

where $\det(I)$ is the determinant of the Fisher Information matrix, and $\operatorname{adj}(I)$ is the adjugated of the matrix.

$$\det(I) = I_{\alpha\alpha}\left(I_{\beta_0\beta_0}I_{\beta_1\beta_1} - I_{\beta_0\beta_1}I_{\beta_1\beta_0}\right) - I_{\alpha\beta_0}\left(I_{\beta_0\alpha}I_{\beta_1\beta_1} - I_{\beta_0\beta_1}I_{\alpha\beta_1}\right) + I_{\alpha\beta_1}\left(I_{\beta_0\alpha}I_{\beta_1\beta_0} - I_{\beta_0\beta_0}I_{\alpha\beta_0}\right)$$

The adjugated matrix is the transpose of the cofactor matrix. Each cofactor C_{ij} is the determinant of the 2×2 submatrix obtained by removing the i-th row and j-th column from I.

$$\operatorname{adj}(I) = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{23} & C_{33} \end{bmatrix}$$

The **Cramér-Rao bounds** provide the minimum possible variances for unbiased estimators. The diagonal elements of the inverse Fisher Information matrix give us these lower bounds:

$$var(\hat{\alpha}) \geq [I^{-1}]_{11}\,, \qquad var\left(\hat{\beta}_0\right) \geq [I^{-1}]_{22}\,, \qquad var\left(\hat{\beta}_1\right) \geq [I^{-1}]_{33}$$

The precision of an estimator is inversely related to its variance: lower variance indicates higher precision. The Cramér-Rao bounds thus set the benchmark for the best achievable precision with unbiased estimators. When an estimator reaches this bound, it is considered efficient.

Problem 10. Estimating Parameters in a Continuous-Time Process

Part a. To derive the log-likelihood function $\ell(\lambda; T_1, T_2, \dots, T_n)$ for a sample of breakdown times, we start with the probability density function (PDF) of the exponential distribution:

$$f_T(t; \lambda) = \lambda e^{-\lambda t}, \qquad t \geq 0$$

Given n independent observations $T_1, T_2, \dots, T_n$, the joint PDF is:

$$f(T_1, T_2, \dots, T_n; \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda T_i}$$

The log-likelihood function $\ell(\lambda; T_1, T_2, \dots, T_n)$ is obtained by taking the natural logarithm of the joint PDF:

$$\ell(\lambda; T_1, T_2, \dots, T_n) = \ln\left(\prod_{i=1}^{n} \lambda e^{-\lambda T_i}\right) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} T_i$$

$$\boldsymbol{\ell(\lambda; T_1, T_2, \dots, T_n) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} T_i}$$

Part b. The **Fisher Information** $I(\lambda)$ is given by the negative expected value of this second derivative:

$$I(\lambda) = E\left[\left(\frac{\partial}{\partial \lambda} \ell(\lambda; T_1, T_2, \dots, T_n)\right)^2 \mid \lambda\right] = E\left[\left(\frac{n}{\lambda} - \sum_{i=1}^{n} T_i\right)^2 \mid \lambda\right]$$

$$= \frac{n^2}{\lambda^2} + E\left[\left(\sum_{i=1}^{n} T_i\right)^2 \mid \lambda\right] - \frac{2n}{\lambda} E\left[\sum_{i=1}^{n} T_i \mid \lambda\right]$$

$$E\left[\sum_{i=1}^{n} T_i \mid \lambda\right] = \sum_{i=1}^{n} E[T_i \mid \lambda] = \sum_{i=1}^{n} \frac{1}{\lambda} = \frac{n}{\lambda}$$

$$E\left[\left(\sum_{i=1}^{n} T_i\right)^2 \mid \lambda\right] = E\left[\sum_{i=1}^{n} T_i \sum_{j=1}^{n} T_j \mid \lambda\right] = E\left[\sum_{i=1}^{n} T_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} T_i T_j\right] = \sum_{i=1}^{n} E[T_i^2] + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} E[T_i] \,.\, E[T_j]$$

$$= \sum_{i=1}^{n} \frac{2}{\lambda^2} + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \frac{1}{\lambda} \,.\, \frac{1}{\lambda} = \frac{2n}{\lambda^2} + \frac{n(n-1)}{\lambda^2} = \frac{n(n+1)}{\lambda^2}$$

$$\text{Fisher Information} \rightarrow I(\lambda) = \frac{n^2}{\lambda^2} + \frac{n(n+1)}{\lambda^2} - \frac{2n^2}{\lambda^2} = \frac{n}{\lambda^2} \rightarrow \boldsymbol{I(\lambda) = \frac{n}{\lambda^2}}$$

The **Cramér-Rao lower bound (CRLB)** provides a lower bound for the variance of any unbiased estimator of a parameter. It is given by the inverse of n times the Fisher Information:

$$\text{CRLB} = \frac{1}{I(\lambda)} = \frac{1}{\frac{n}{\lambda^2}} = \frac{\lambda^2}{n} \rightarrow \textbf{CRLB} = \boldsymbol{\frac{\lambda^2}{n}}$$

Part c. To find the **maximum likelihood estimator (MLE)**, we maximize this log-likelihood function with respect to λ. We start by taking the first derivative and setting it to zero:

$$\frac{\partial}{\partial\lambda}\ell(\lambda; T_1, T_2, \ldots, T_n) = \frac{\partial}{\partial\lambda}\left(n\ln(\lambda) - \lambda\sum_{i=1}^{n} T_i\right) = \frac{n}{\lambda} - \sum_{i=1}^{n} T_i = 0 \rightarrow \hat{\lambda}_{MLE} = \frac{n}{\sum_{i=1}^{n} T_i} = \frac{1}{\overline{T_n}} \rightarrow \boldsymbol{\hat{\lambda}_{MLE} = \frac{1}{\overline{T_n}}}$$

Variance of $\hat{\lambda}_{MLE}$: To compute the variance of $\hat{\lambda}_{MLE}$, note that $S_n = \sum_{i=1}^{n} T_i$, where each T_i is exponentially distributed with mean $\frac{1}{\lambda}$. The sum of n independent exponential random variables with rate λ follows a Gamma distribution with shape parameter n and rate λ, denoted $\text{Gamma}(n, \lambda)$.

The mean and variance of S_n are:

$$E[S_n] = \frac{n}{\lambda}, \qquad var(S_n) = \frac{n}{\lambda^2}$$

The Delta Method for Variance Approximation:

$$\text{var}(g(X)) \approx \left(g'(E[X])\right)^2 \text{var}(X)$$

Since $\hat{\lambda}_{MLE} = \frac{n}{S}$, we use the delta method to approximate the variance of $\hat{\lambda}_{MLE}$:

$$\hat{\lambda}_{MLE} = \frac{1}{\overline{T_n}} = \frac{n}{S_n} = g(S_n)$$

$$\text{var}(\hat{\lambda}_{MLE}) \approx \left(\frac{\partial\left(\frac{n}{S_n}\right)}{\partial S_n}\right)^2 \text{var}(S_n) = \left(-\frac{n}{S_n^2}\right)^2 \text{var}(S_n) = \frac{n^2}{S_n^4} \cdot \frac{n}{\lambda^2} = \frac{n^3}{\lambda^2 S_n^4}$$

Using $S \approx E[S] = \frac{n}{\lambda}$ for large n, this simplifies to:

$$\text{var}(\hat{\lambda}_{MLE}) \approx \frac{n^3}{\lambda^2\left(\frac{n}{\lambda}\right)^4} = \frac{\lambda^2}{n}$$

The CRLB for the variance of any unbiased estimator of λ is:

$$\boldsymbol{\text{var}(\hat{\lambda}_{MLE}) \approx \frac{\lambda^2}{n}}$$

Since the variance of $\hat{\lambda}_{MLE}$ matches $\frac{\lambda^2}{n}$, the MLE achieves the CRLB, confirming it is efficient.

$$\text{var}(\hat{\lambda}_{MLE}) \approx \frac{\lambda^2}{n} \geq \text{CRLB} = \frac{\lambda^2}{n}$$

As n increases, the variance decreases proportionally to $\frac{1}{n}$:

$$\text{var}(\hat{\lambda}_{MLE}) \propto \frac{1}{n}$$

This inverse relationship implies that the precision of $\hat{\lambda}_{MLE}$ improves as the sample size grows. For large n, the estimator becomes increasingly concentrated around the true value of λ, reflecting the efficiency of the MLE.

Problem 11. Analyzing Investment Returns Using MGFs

Part a. Moment-Generating Function (MGF)

For a normal distribution with mean μ and variance σ^2, the MGF is:

$$M_X(t) = e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}$$

Given the parameters for the stock's annual return X, where $\mu = 0.08$ and $\sigma^2 = 0.04$:

$$M_X(t) = e^{0.08t} e^{\frac{1}{2}0.04t^2}$$

Therefore, the moment-generating function (MGF) of the random variable X is:

$$\boldsymbol{M_X(t) = e^{0.08t} e^{0.02t^2}}$$

Part b. Mean and Variance from the MGF

The expected value $E[X]$ can be found by taking the first derivative of the MGF with respect to t and then evaluating it at $t = 0$:

$$\boldsymbol{E[X]} = \mu_X = \left[\frac{\partial M_X(t)}{\partial t}\right]_{t=0} = \left[\frac{\partial}{\partial t}\left(e^{0.08t} e^{0.02t^2}\right)\right]_{t=0} = \left[(0.08 + 0.04t)\left(e^{0.08t} e^{0.02t^2}\right)\right]_{t=0} = \mathbf{0.08}$$

The variance $var(X)$ is found using the second derivative of the MGF. First, let's find the second derivative of the MGF and then evaluate it at $t = 0$:

$$\begin{aligned}
\boldsymbol{var(X)} = \sigma_X^2 &= \left[\frac{\partial^2 M_X(t)}{\partial t^2}\right]_{t=0} - (E[X])^2 = \left[\frac{\partial}{\partial t}\left((0.08 + 0.04t)\left(e^{0.08t} e^{0.02t^2}\right)\right)\right]_{t=0} - 0.0064 \\
&= \left[\frac{\partial}{\partial t}\left((0.08 + 0.04t)M_x(t)\right)\right]_{t=0} - 0.0064 \\
&= \left[0.04\left(e^{0.08t} e^{0.02t^2}\right) + (0.08 + 0.04t)^2 e^{0.08t} e^{0.02t^2}\right]_{t=0} - 0.0064 \\
&= 0.04 + 0.0064 - 0.0064 = \mathbf{0.04}
\end{aligned}$$

Part c. Calculate Higher Moments

The third moment can be found by taking the third derivative of the MGF with respect to t and evaluating it at $t = 0$:

$$\begin{aligned}
\boldsymbol{E[X^3]} &= \left[\frac{\partial^3 M_X(t)}{\partial t^3}\right]_{t=0} = \left[\frac{\partial^2}{\partial t^2}\left((0.08 + 0.04t)M_x(t)\right)\right]_{t=0} = \left[\frac{\partial}{\partial t}\left(\left((0.08 + 0.04t)^2 + 0.04\right)M_x(t)\right)\right]_{t=0} \\
&= \left[M_X(t)\left((0.08 + 0.04t)^3 + 0.12(0.08 + 0.04t)\right)\right]_{t=0} = 0.0101
\end{aligned}$$

The skewness of a distribution is defined as:

$$\text{Skewness}(X) = \frac{E[(X-\mu)^3]}{\left(var(X)\right)^{\frac{3}{2}}}$$

$$\begin{aligned}
E[(X-\mu)^3] &= E[X^3 - 3\mu X^2 + 3\mu^2 X - \mu^3] = E[X^3] - 3\mu E[X^2] + 3\mu^2 E[X] - \mu^3 \\
&= E[X^3] - 3\mu(\sigma^2 + \mu^2) + 3\mu^3 - \mu^3 = E[X^3] - 3\mu\sigma^2 - \mu^3 \\
&= 0.0101 - 3 \times 0.08 \times 0.04 - 0.08^3 = 0.0101 - 0.0101 = 0
\end{aligned}$$

$$\rightarrow \ \textbf{Skewness}(\boldsymbol{X}) = \mathbf{0}$$

Also, we have: $E[X^3] - 3\mu\sigma^2 - \mu^3 = 0 \rightarrow E[X^3] = 3\mu\sigma^2 + \mu^3$

Problem 12. Fisher Information and Bayesian Inference (Optional)

Part a. Given the Gamma distribution:

$$f(x;\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$$

The MLE for β is given by: $\hat{\beta}=\frac{\alpha}{\bar{X}}$

knowing that $\alpha=5$, we aim to **estimate $\boldsymbol{\beta}$**.

- $\boldsymbol{n=200}$: $\quad \boldsymbol{\hat{\beta}=\frac{\alpha}{\bar{X}}=\frac{5}{\bar{X}}=\frac{1000}{\sum_{i=1}^{200}X_i}}$
- $\boldsymbol{n=1000}$: $\quad \boldsymbol{\hat{\beta}=\frac{\alpha}{\bar{X}}=\frac{5}{\bar{X}}=\frac{5000}{\sum_{i=1}^{1000}X_i}}$

The Fisher Information for β is given by: $I(\beta)=\frac{n\alpha}{\beta^2}$

- $\boldsymbol{n=200}$: $\quad \boldsymbol{I(\beta)=\frac{1000}{\beta^2}}$
- $\boldsymbol{n=1000}$: $\quad \boldsymbol{I(\beta)=\frac{5000}{\beta^2}}$

Part b.

- **Prior Distribution:** Assume a Gamma prior for β with hyperparameters $\alpha_0=2$ and $\beta_0=1$:

$$\pi(\beta)=\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\beta^{\alpha_0-1}\,e^{\beta_0\beta}$$

- **Posterior Distribution:** Using Bayesian updating for the two sample sizes:

$$\alpha_n=\alpha_0+n\alpha\,,\qquad \beta_n=\beta_0+n\beta$$
$$\pi(\beta\mid X)\sim\mathrm{Gamma}(\alpha_n,\beta_n)$$

Fisher Information of the Posterior: For each sample size, compute the Fisher Information of the posterior distribution:

$$I_{posterior}(\beta)=\frac{\alpha_n}{\beta_n^2}=\frac{\alpha_0+n\alpha}{(\beta_0+n\beta)^2}$$

Part c. For the MLE, the 95% confidence interval for β is calculated as:

$$\mathrm{CI}_{0.95}=\left(\hat{\beta}-\frac{1.96\hat{\beta}}{\sqrt{n}}\,,\hat{\beta}+\frac{1.96\hat{\beta}}{\sqrt{n}}\right)$$

For the posterior distribution, the 95% credible interval is derived from the quantiles of the Gamma posterior distribution. Specifically:

$$\mathrm{CI}_{0.95}=\left(q_{\frac{\alpha}{2}}\,,q_{1-\frac{\alpha}{2}}\right)$$

where $q_{\frac{\alpha}{2}}$ and $q_{1-\frac{\alpha}{2}}$ are the $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ quantiles of the Gamma distribution, respectively.

Problem 13. Analyzing Investment Returns Using MGF

Part a. Moment-Generating Function (MGF)

- *Analytical Task:*

For a normal distribution with mean μ and variance σ^2, the MGF is:

$$M_X(t) = e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}$$

Given the parameters for the stock's annual return X, where $\mu = 0.08$ and $\sigma^2 = 0.04$:

$$M_X(t) = e^{0.08t} e^{\frac{1}{2}0.04t^2}$$

Therefore, the moment-generating function (MGF) of the random variable X is:

$$\boldsymbol{M_X(t) = e^{0.08t} e^{0.02t^2}}$$

- *Computation Task:*

To simulate some random data with the given Gaussian distribution, we generate the data and then show the MGF in the range -10 to 10 using the requested function. As seen in figure 13.1, the two charts overlap, so the theoretical calculations have been done correctly.

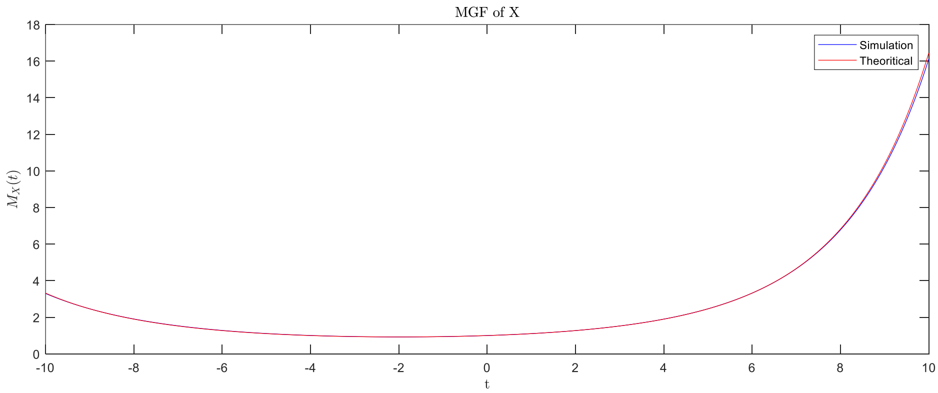


Figure 13.1. Plotting the MGF of X

Part b. Mean and Variance from the MGF

- *Analytical Task:*

The expected value $E[X]$ can be found by taking the first derivative of the MGF with respect to t and then evaluating it at $t = 0$:

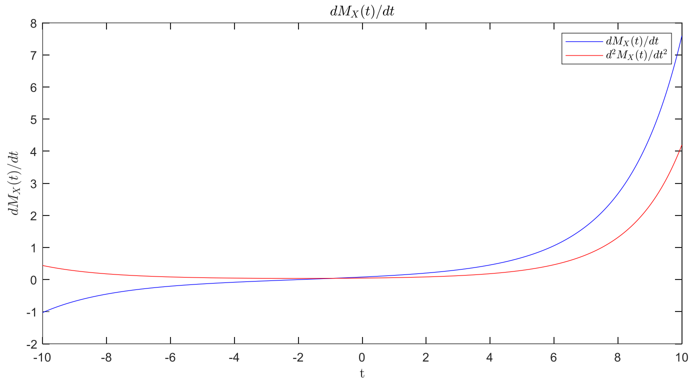
$$\boldsymbol{E[X]} = \mu_X = \left[\frac{\partial M_X(t)}{\partial t}\right]_{t=0} = \left[\frac{\partial}{\partial t}\left(e^{0.08t} e^{0.02t^2}\right)\right]_{t=0} = \left[(0.08 + 0.04t)\left(e^{0.08t} e^{0.02t^2}\right)\right]_{t=0} = \mathbf{0.08}$$

The variance $var(X)$ is found using the second derivative of the MGF. First, let's find the second derivative of the MGF and then evaluate it at $t = 0$:

$$\begin{aligned}\boldsymbol{var(X)} = \sigma_X^2 &= \left[\frac{\partial^2 M_X(t)}{\partial t^2}\right]_{t=0} - (E[X])^2 = \left[\frac{\partial}{\partial t}\left((0.08 + 0.04t)\left(e^{0.08t} e^{0.02t^2}\right)\right)\right]_{t=0} - 0.0064 \\ &= \left[\frac{\partial}{\partial t}\left((0.08 + 0.04t) M_x(t)\right)\right]_{t=0} - 0.0064 \\ &= \left[0.04\left(e^{0.08t} e^{0.02t^2}\right) + (0.08 + 0.04t)^2 e^{0.08t} e^{0.02t^2}\right]_{t=0} - 0.0064 \\ &= 0.04 + 0.0064 - 0.0064 = \mathbf{0.04}\end{aligned}$$

- *Computation Task:*

To calculate the first and second moments, we use the *diff* command in MATLAB to derive the MGF. Figure 13.2 shows the mean and variance as well as the shape of the first and second moments. As we can see, the mean and variance match the theoretical values.




```
index0 = find(t == 0);
E_X = dM_sim(index0 + 1)

E_X = 0.0798

E_X2 = d2M_sim(index0 + 2);
var_X = E_X2 - E_X^2

var_X = 0.0406
```

Figure 13.2. Plotting the MGF of X

Part c. Calculate Higher Moments

- *Analytical Task:*

The third moment can be found by taking the third derivative of the MGF with respect to t and evaluating it at $t = 0$:

$$\boldsymbol{E}[\boldsymbol{X}^3] = \left[\frac{\partial^3 M_X(t)}{\partial t^3}\right]_{t=0} = \left[\frac{\partial^2}{\partial t^2}\big((0.08 + 0.04t)M_x(t)\big)\right]_{t=0} = \left[\frac{\partial}{\partial t}\big(((0.08 + 0.04t)^2 + 0.04)M_x(t)\big)\right]_{t=0}$$
$$= \left[M_X(t)\big((0.08 + 0.04t)^3 + 0.12(0.08 + 0.04t)\big)\right]_{t=0} = 0.0101$$

The skewness of a distribution is defined as:

$$\text{Skewness}(X) = \frac{E[(X-\mu)^3]}{\big(var(X)\big)^{\frac{3}{2}}}$$

$$E[(X-\mu)^3] = E[X^3 - 3\mu X^2 + 3\mu^2 X - \mu^3] = E[X^3] - 3\mu E[X^2] + 3\mu^2 E[X] - \mu^3$$
$$= E[X^3] - 3\mu(\sigma^2 + \mu^2) + 3\mu^3 - \mu^3 = E[X^3] - 3\mu\sigma^2 - \mu^3$$
$$= 0.0101 - 3 \times 0.08 \times 0.04 - 0.08^3 = 0.0101 - 0.0101 = 0$$

$$\rightarrow \textbf{Skewness}(\boldsymbol{X}) = \mathbf{0}$$

Also, we have: $E[X^3] - 3\mu\sigma^2 - \mu^3 = 0 \rightarrow E[X^3] = 3\mu\sigma^2 + \mu^3$

- *Computation Task:*

In this section, similar to the previous one, we obtain the third moment. As observed, this value is almost equal to the theoretical value. The reason for the difference between the theoretical and practical skewness values is that the variance is small, which causes a slight error in the division related to skewness, reducing the accuracy.

```
d3M_sim = diff(d2M_sim) / dt;
E_X3 = d3M_sim(index0 + 3)

E_X3 = 0.0102

skewness_X = (E_X3 - 3 * E_X * E_X2 + 3 * (E_X)^2 * E_X2 - (E_X)^3) / (var_X)^1.5

skewness_X = -0.0821
```

Figure 13.3. Plotting the MGF of X

Problem 14. Plants vs. Enemies II (Optional)

The triangular farm is defined by the line $X + Y = 1$, bounded by the points $(0,0)$, $(1,0)$, and $(0,1)$.

1- The enemy's movement along the x-axis follows a uniform distribution $U[0,1]$. If the farmer farms along the x-axis at a point x, the enemy's movement will uniformly cover $[0, x]$, leading to a successful farming area of $1 - x$.
2- The enemy's movement along the y-axis follows a uniform distribution $U[0,1 - X]$. If the farmer farms along the y-axis at a point y, the enemy's movement will uniformly cover $[0, y]$, leading to a successful farming area of $1 - y$.

We need to compare the expected values of the successful farming areas for both cases.

$$E[X] = \int_0^1 (1-x)dx = \left[x - \frac{x^2}{2}\right]_0^1 = \left(1 - \frac{1}{2}\right) - (0-0) = \frac{1}{2}$$

$$E[Y] = \int_0^1 (\int_0^{1-x} (1-y)dy)dx = \int_0^1 \left[y - \frac{y^2}{2}\right]_0^{1-x} dx = \int_0^1 \left(1 - x - \frac{1}{2}(1-x)^2\right) dx = \int_0^1 \left(\frac{1}{2} - \frac{x^2}{2}\right) dx$$
$$= \left[\frac{x}{2} - \frac{x^3}{6}\right]_0^1 = \frac{1}{3}$$

The expected successful farming area along the x-axis is $\frac{1}{2}$, while the expected successful farming area along the y-axis is $\frac{1}{3}$.

Thus, it is better for the farmer to farm along the **x-axis** to maximize the expected successful farming area.