

Exercise 1. [5.5 point(s)] Let Z_1, \dots, Z_n be independent standard normal random variables. For $i = 1, \dots, n$, denote $X_i = Z_1 + \dots + Z_i$.

1. [1 point(s)] Let $i \in \{1, \dots, n\}$. What is the distribution of X_i ?
2. [1 point(s)] Let $i, j \in \{1, \dots, n\}$. What is the covariance between X_i and X_j ?
3. [1 point(s)] What is the joint distribution of (X_1, \dots, X_n) ?
4. Let $k \in \{1, \dots, n\}$ and $x \in \mathbb{R}$.
 - (a) [1 point(s)] What is the distribution of X_n knowing that $X_k = x$?
 - (b) [1.5 point(s)] What is the distribution of X_k knowing that $X_n = x$?

Exercise 2. [3.5 point(s)] A car rental wants to determine which of its cars have been used in the city only. In order to do this, it installs a device that records the number of hours the car is driven on each one of its gears. We know that a car used in the city spends on average 10% of its time in the neutral position, 5% on the first gear, 30% on the second gear, 30% on the third gear, 20% on the fourth gear, and 5% on the fifth gear. The agency decides to run a test based on the Pearson statistic in order to decide whether a car was used in the city only. It chooses the significance level $\alpha = 5\%$.

1. A first car was driven $210h$ in the neutral position, $94h$ in the first gear, $564h$ in the second gear, $630h$ in the third gear, $390h$ in the fourth gear and $112h$ in the fifth gear. Was it used in the city only?
2. A second car was driven $220h$ in the neutral position, $80h$ in the first gear, $340h$ in the second gear, $600h$ in the third gear, $480h$ in the fourth gear and $280h$ in the fifth gear. Was it used in the city only?

Exercise 3. [5.5 point(s)] Random variable X follows Erlang distribution $f_X(x) = \frac{1}{24} \lambda^5 x^4 e^{-\lambda x} I(x > 0)$ where $\lambda > 0$ is the parameter of the distribution. Suppose that $n = 2000$ i.i.d. samples $x_1, \dots, x_n > 0$ have been observed from this distribution. Moreover, we know that the average of the observed samples is 2, i.e., $\frac{x_1 + x_2 + \dots + x_n}{n} = 2$. In this question, you need to report numerical values as well as providing a complete solution.

1. [3 point(s)] Compute the maximum likelihood estimator for λ .
2. We now want to build a confidence interval for λ .
 - (a) [0.75 point(s)] Compute the observed information $J(\lambda) = -\frac{d^2 \ell(\lambda)}{d\lambda^2}$ (where ℓ represents the log-likelihood function). What is $J(\hat{\lambda}_{MLE})$ where $\hat{\lambda}_{MLE}$ is the maximum likelihood estimator?
 - (b) [1.75 point(s)] A result from the lecture notes of the course states that

$$J(\hat{\lambda})^{1/2} (\hat{\lambda} - \lambda) \xrightarrow[n \rightarrow \infty]{D} Z,$$

where Z is a standard normal random variable.

From this result, deduce an approximate pivot. Use this pivot to derive an equi-tailed 95% confidence interval for λ .

Exercise 4. [3 point(s)] Consider a coin that falls on heads with probability p . In an experiment, the coin is tossed and the first outcome is noted. Then the tossing is continued until the outcome is the complement of the first outcome. For example, if the first coin is heads, we continue tossing the coin until we see tails. Let N be the number of tosses other than the first toss. For example, assume the first toss is tails. We then repeatedly throw the coin until we fall on heads, for instance obtaining the sequence tails, tails, heads. This outcome gives $N = 3$. (Again, we do not count the very first toss.) Find the probability mass function of N and find $\mathbb{E}[N]$.

Exercise 5. [5.5 point(s)] Let X, Y be two independent standard normal random variables. Denote $U = X/Y$. Note that U is well-defined as the event $\{Y = 0\}$ has probability zero.

1. [0.75 point(s)] What is the joint probability density function of (X, Y) ?
2. [2.25 point(s)] What is the distribution of $\frac{1}{2}(X + Y)^2 + \frac{1}{2}(X - Y)^2$?
3. [1.5 point(s)] What is the joint probability density function of (U, Y) ?
4. [1 point(s)] What is the probability density function of U ?

Exercise 6. [4.5 point(s)] m people are in a room and each of them selects a random integer uniformly in $\{1, 2, \dots, n\}$, independently of others. Denote p the probability that at least two people have selected the same number.

- 1.a. [1 point(s)] For $1 \leq i < j \leq m$, define random variable $Y_{i,j}$ to be 1 if person i and j chose the same integer and 0 otherwise. Also define $Y = \sum_{i < j} Y_{i,j}$. Compute $\mathbb{E}[Y]$.
- 1.b. [1 point(s)] Using Markov's inequality, conclude that if $m \leq \sqrt{n}$ then $p < 0.5$.
2. [2.5 point(s)] Show that if $m \geq \sqrt{2n} + 1$ then $p > 0.5$.

Exercise 7. [3 point(s)] Let $n \in \mathbb{N}$. Show that

$$2^{-n} \sum_{\substack{k \in \{0, \dots, n\} \\ \text{such that } |k - \frac{n}{2}| \leq \frac{\sqrt{n}}{2}}} \binom{n}{k} \xrightarrow{n \rightarrow \infty} \int_{-1}^1 \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du.$$