Probability and Statistics
FS 2017
Second Session Exam
09.02.2018
Time Limit: 180 Minutes

Name: ____________________

Student ID: ____________________

This exam contains 19 pages (including this cover page) and 10 questions. A Formulae sheet is provided with the exam.

Please justify all your steps carefully. Otherwise no points will be given.

Grade Table (for grading use only, please leave empty)

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

Informations. Read this carefully.

- Please justify all your statements carefully. Explain the steps of your reasoning. Otherwise no points will be given.
- You are expected to write full sentences when giving your answer.
- DO NOT WRITE with red or green pens. DO NOT WRITE with a pencil.
- Your answers should be *readable*.
- Write your name on all the sheets you intend to hand in before the end of the exam.

GOOD LUCK

1. (10 points) **Conditional Probabilities**

   (a) (4 points) Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $A, B \in \mathcal{A}$ with $\mathbb{P}(A) = 0.6$ and $\mathbb{P}(B) = 0.7$. *Show that* $\mathbb{P}(B|A) \geq 0.5$.

   (b) (6 points) Suppose we are given three different dice. The first die is fair, i.e. the probability to obtain a six is $\frac{1}{6}$. The second and the third die are biased. The probability to obtain a six with the second die is $\frac{1}{2}$ and the probability to obtain a six with the third die is 1. Suppose that one of the three dice is chosen at random and tossed. Moreover, suppose that we get a six. *Calculate the probability that the fair die was chosen.*

Solution

(a) We calculate

$$\mathbb{P}(A^c \cup B^c) \leq \mathbb{P}(A^c) + \mathbb{P}(B^c) = 0.4 + 0.3 = 0.7,$$

which leads to

$$\mathbb{P}(A \cap B) = 1 - \mathbb{P}(A^c \cup B^c) \geq 1 - 0.7 = 0.3.$$

Therefore we get

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \geq \frac{0.3}{0.6} = 0.5.$$

(b) We define the following four events:
$A :=$ first die was chosen,
$B :=$ second die was chosen,
$C :=$ third die was chosen,
"6" $:=$ we obtained a six,
i.e. we have to calculate $\mathbb{P}(A|\text{"6"})$. We know from the problem formulation that

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{3}$$

and that

$$\mathbb{P}(\text{"6"}|A) = \frac{1}{6}, \quad \mathbb{P}(\text{"6"}|B) = \frac{1}{2} \quad \text{and} \quad \mathbb{P}(\text{"6"}|C) = 1.$$

We apply Bayes Theorem to get

$$\begin{aligned}
\mathbb{P}(A|\text{"6"}) &= \frac{\mathbb{P}(\text{"6"}|A) \cdot \mathbb{P}(A)}{\mathbb{P}(\text{"6"}|A) \cdot \mathbb{P}(A) + \mathbb{P}(\text{"6"}|B) \cdot \mathbb{P}(B) + \mathbb{P}(\text{"6"}|C) \cdot \mathbb{P}(C)} \\
&= \frac{\frac{1}{6} \cdot \frac{1}{3}}{\frac{1}{6} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} \\
&= \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{2} + 1} \\
&= \frac{1}{10}.
\end{aligned}$$

2. (10 points) **Density, Expectation, Variance and Covariance**

Let X be a real-valued random variable with density

$$f(x) = \begin{cases} c(4-x^2) & \text{if } 0 \le x \le 2, \\ 0 & \text{else,} \end{cases}$$

for some constant $c > 0$.

(a) (2 points) *Calculate* c.

(b) (2 points) *Calculate* $\mathbb{E}[X]$.

(c) (2 points) *Calculate* $\mathrm{Var}(X)$.

Let Y be another real-valued random variable with $\mathbb{E}[Y] = \frac{5}{4}$, $\mathrm{Var}(Y) = \frac{101}{80}$ and $\mathrm{Cov}(X,Y) = \frac{1}{4}$.

(d) (2 points) *Calculate* $\mathbb{E}[XY]$.

(e) (2 points) *Calculate* $\mathrm{Var}(X+Y)$.

Solution

(a) We calculate

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^2 c(4-x^2)\,dx = c\left(8 - \frac{8}{3}\right) = c\,\frac{16}{3}.$$

Since f is the density of a random variable, it has to integrate to 1. Hence, we get

$$c = \frac{3}{16}.$$

(b) To get $\mathbb{E}[X]$, we compute

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^2 x\frac{3}{16}(4-x^2)\,dx = \frac{3}{16}\int_0^2 4x - x^3\,dx = \frac{3}{16}(8-4) = \frac{3}{4}.$$

(c) In order to calculate $\mathrm{Var}(X)$, we need

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^2 x^2\frac{3}{16}(4-x^2)\,dx = \frac{3}{16}\int_0^2 4x^2 - x^4\,dx = \frac{3}{16}\left(\frac{32}{3} - \frac{32}{5}\right) = \frac{4}{5}.$$

Hence, we get

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{4}{5} - \left(\frac{3}{4}\right)^2 = \frac{64}{80} - \frac{45}{80} = \frac{19}{80}.$$

(d) We calculate

$$\mathbb{E}[XY] = \mathrm{Cov}(X,Y) + \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{4} + \frac{3}{4}\cdot\frac{5}{4} = \frac{19}{16}.$$

(e) We calculate

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) = \frac{19}{80} + \frac{101}{80} + 2\cdot\frac{1}{4} = 2.$$

3. (10 points) **Chebyshev Inequality**

(a) (4 points) Let $g : \mathbb{R} \to [0, \infty)$ be an increasing function. Let $c \in \mathbb{R}$ such that $g(c) > 0$. Let X be a real-valued random variable. *Show that*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[g(X)]}{g(c)}.$$

Note that you can NOT use that X has a probability mass function or a density. Hint: Use that the probability of an event can be written as the expectation of the indicator function of this event.

(b) (2 points) Let X be a real-valued random variable and $c > 0$ a constant. *Use the result in (a) to show that*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq c) \leq \frac{\operatorname{Var}(X)}{c^2}.$$

Now let $X_1, \dots, X_{50}$ be independent, identically Poisson distributed random variables with parameter $\lambda = 2$ and define $S = \frac{1}{50} \sum_{i=1}^{50} X_i$.

(c) (4 points) *Using the result in (b), give the smallest possible value of $c > 0$ such that*

$$\mathbb{P}(|S - \mathbb{E}[S]| \geq c) \leq 0.01.$$

Solution

(a) Since g is positive, increasing and $g(c) > 0$, we get

$$\mathbf{1}_{\{X \geq c\}} \leq \frac{g(X)}{g(c)} \quad \text{almost surely.}$$

Hence we can conclude

$$\mathbb{P}(X \geq c) = \mathbb{E}[\mathbf{1}_{\{X \geq c\}}] \leq \mathbb{E}\left[\frac{g(X)}{g(c)}\right] = \frac{\mathbb{E}[g(X)]}{g(c)}.$$

(b) We define $Y := |X - \mathbb{E}[X]|$ and $g(x) := x^2$ as a function from $[0, \infty)$ to $[0, \infty)$. Then g is an increasing function with $g(x) > 0$ for all $x > 0$ and

$$\mathbb{E}[g(Y)] = \mathbb{E}[(X - \mathbb{E}[X])^2)] = \operatorname{Var}(X).$$

Hence we get

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq c) = \mathbb{P}(Y \geq c) \leq \frac{\mathbb{E}[g(Y)]}{g(c)} = \frac{\operatorname{Var}(X)}{c^2},$$

where for the inequality we used the result in (a).

(c) Let $i \in \{1, \dots, 50\}$. Since $X_i \sim \mathrm{Poi}(2)$, we have $\mathrm{Var}(X_i) = 2$. Moreover, since $X_1, \dots, X_{50}$ are independent, we have

$$\mathrm{Var}(S) = \mathrm{Var}\left(\frac{1}{50}\sum_{i=1}^{50} X_i\right) = \left(\frac{1}{50}\right)^2 \cdot \sum_{i=1}^{50} \mathrm{Var}(X_i) = \frac{2}{50} = \frac{1}{25}.$$

Using the result in (b), we get

$$\mathbb{P}(|S - \mathbb{E}[S]| \geq c) \leq \frac{\mathrm{Var}(S)}{c^2} = \frac{1}{25c^2}.$$

Now

$$\frac{1}{25c^2} \leq 0.01 \qquad \Longleftrightarrow \qquad c^2 \geq 4,$$

from which we can conclude that $c = 2$.

4. (10 points) **Convergence in Probability and Almost Sure Convergence**

   (a) (3 points) Let $(X_n)_{n\geq 1}$ be independent, Bernoulli distributed random variables with

$$\mathbb{P}(X_n = 1) = \frac{1}{n} \quad \text{and} \quad \mathbb{P}(X_n = 0) = 1 - \frac{1}{n}, \qquad \text{for all } n \geq 1.$$

   Show that X_n converges to 0 in probability, as $n \to \infty$.

   (b) (7 points) Let $(Y_n)_{n\geq 1}$ be independent, identically uniformly distributed random variables on $[0,1]$. Define $M_n := \min\{Y_1, \ldots, Y_n\}$ for all $n \geq 1$. *Show that M_n converges to 0 almost surely, as $n \to \infty$.*
   Hint: Use the first Borel-Cantelli Lemma to show that for an arbitrary $y \in (0,1)$ we have $\lim\limits_{n\to\infty} M_n \leq y$ almost surely.

Solution

(a) For $\epsilon \geq 1$, we get

$$\mathbb{P}(|X_n - 0| > \epsilon) = \mathbb{P}(X_n > \epsilon) = 0, \qquad \text{for all } n \geq 1.$$

For $0 < \epsilon < 1$, we get

$$\mathbb{P}(|X_n - 0| > \epsilon) = \mathbb{P}(X_n > \epsilon) = \mathbb{P}(X_n = 1) = \frac{1}{n} \xrightarrow{n\to\infty} 0.$$

(b) Let $y \in (0,1)$ and $n \geq 1$. Then, using the i.i.d.-property of $Y_1, \ldots, Y_n$ and the definition of the distribution function of a uniform distribution on $[0,1]$, we calculate

$$\begin{aligned}\mathbb{P}(M_n > y) &= \mathbb{P}(\min\{Y_1, \ldots, Y_n\} > y) = \mathbb{P}(Y_1 > y, \ldots, Y_n > y) = \prod_{i=1}^{n} \mathbb{P}(Y_i > y) \\ &= \mathbb{P}(Y_1 > y)^n = (1-y)^n .\end{aligned}$$

Using this result, we get

$$\sum_{n=1}^{\infty} \mathbb{P}(M_n > y) = \sum_{n=1}^{\infty} (1-y)^n = \frac{1-y}{1-(1-y)} = \frac{1}{y} - 1 < \infty.$$

Due to the first Borel-Cantelli Lemma, we get $\mathbb{P}(M_n > y \text{ for infinitely many } n) = 0$. This implies that $\lim_{n\to\infty} M_n \leq y$ almost surely. Since y was chosen arbitrarily in $(0,1)$, we can conclude that M_n converges to 0 almost surely, as $n \to \infty$.

5. (10 points) **Random Vector and Conditional Distribution**

Let X and Y be two independent, exponentially distributed random variables with parameter $\lambda = 1$. Let us define $T = X + Y$.

(a) (1 point) *What is the joint density of the random vector (X, Y)?*

(b) (5 points) *Calculate the joint density of the random vector (X, T).*

(c) (2 points) *Use the result obtained in (b) to calculate the marginal density of T.*

(d) (2 points) *Calculate the conditional density of X given $T = 1$.*

Solution

(a) Since X and Y both have an Exponential(1)-distribution, their corresponding densities f_X and f_Y are

$$f_X(x) = f_Y(x) = \mathrm{e}^{-x} \cdot \mathbf{1}_{\{x \geq 0\}}.$$

Moreover, since X and Y are independent, the joint density $f_{X,Y}$ of the random vector (X, Y) is given by

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) = \mathrm{e}^{-(x+y)} \cdot \mathbf{1}_{\{x,y \geq 0\}}.$$

(b) We define

$$B := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then

$$\begin{pmatrix} X \\ X+Y \end{pmatrix} = B \cdot \begin{pmatrix} X \\ Y \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad \det(B) = 1.$$

By the Transformation Theorem, we get for the joint density $f_{X,T}$ of the random vector (X, T)

$$\begin{aligned} f_{X,T}(x,t) &= \frac{1}{|\det(B)|} \cdot f_{X,Y}\left(B^{-1} \cdot \begin{pmatrix} x \\ t \end{pmatrix}\right) \\ &= f_{X,Y}(x, t-x) \\ &= \mathrm{e}^{-(x+t-x)} \cdot \mathbf{1}_{\{x, t-x \geq 0\}} \\ &= \mathrm{e}^{-t} \cdot \mathbf{1}_{\{t \geq x \geq 0\}}. \end{aligned}$$

(c) We calculate the marginal density f_T of T using the joint density $f_{X,T}$ found in (b):

$$f_T(t) = \int_{-\infty}^{\infty} f_{X,T}(x,t)\,dx = \int_{-\infty}^{\infty} \mathrm{e}^{-t} \cdot \mathbf{1}_{\{t \geq x \geq 0\}}\,dx = \int_0^t \mathrm{e}^{-t} \cdot \mathbf{1}_{\{t \geq 0\}}\,dx = t\,\mathrm{e}^{-t} \cdot \mathbf{1}_{\{t \geq 0\}}.$$

(d) We calculate the conditional density $f_{X \,|\, T=1}$ of $X \,|\, T = 1$ using the joint density $f_{X,T}$ found in (b) and the marginal density f_T found in (c):

$$f_{X \,|\, T=1}(x \,|\, 1) = \frac{f_{X,T}(x,1)}{f_T(1)} = \frac{\mathrm{e}^{-1} \cdot \mathbf{1}_{\{1 \geq x \geq 0\}}}{1 \cdot \mathrm{e}^{-1}} = \mathbf{1}_{\{1 \geq x \geq 0\}}.$$

6. (10 points) **Joint Distribution and Jensen's Inequality**

Peter has two possibilities to go to work: Either he can walk the short distance or he can take the bus. The bus stop is on his way to the office. Every morning Peter arrives at a random time between 7:10 and 7:20 at the bus stop. Likewise, also the bus arrives at the stop at a random time between 7:10 and 7:20. Let $X \sim \mathrm{Uni}([10,20])$ model the arrival time of Peter at the bus stop in minutes after 7:00 and $Y \sim \mathrm{Uni}([10,20])$ model the arrival time of the bus at the bus stop in minutes after 7:00. Moreover, we assume that the arrival time of Peter and the arrival time of the bus at the bus stop are independent, i.e. that X and Y are independent.

(a) (5 points) Suppose that when Peter arrives at the bus stop, he waits for at most five minutes: If the bus arrives within these five minutes, he takes the bus, otherwise he walks to the office. *Calculate the probability that Peter takes the bus.*

(b) (5 points) *Show that* $\mathbb{E}\left[\frac{X}{Y}\right] > 1$ *without calculating it. Can we conclude that, on average, Peter arrives at the bus stop later than the bus?*

Solution

(a) Since X and Y are independent, uniformly on $[10,20]$ distributed random variables, the joint density $f_{X,Y}$ of (X,Y) is given by

$$f_{X,Y}(x,y) = \frac{1}{10} \cdot \mathbf{1}_{\{x \in [10,20]\}} \cdot \frac{1}{10} \cdot \mathbf{1}_{\{y \in [10,20]\}} = \frac{1}{100} \cdot \mathbf{1}_{\{x,y \in [10,20]\}}.$$

Peter takes the bus if and only if $X \leq Y \leq X+5$. Thus, we calculate

$$\begin{aligned}
\mathbb{P}(\text{"Peter takes the bus"}) &= \mathbb{P}(X \leq Y \leq X+5) \\
&= \int_{10}^{20} \int_{x}^{x+5} f_{X,Y}(x,y)\, dydx \\
&= \int_{10}^{15} \int_{x}^{x+5} \frac{1}{100}\, dydx + \int_{15}^{20} \int_{x}^{20} \frac{1}{100}\, dydx \\
&= \frac{1}{100} \int_{10}^{15} (x+5-x)\, dx + \frac{1}{100} \int_{15}^{20} (20-x)\, dx \\
&= \frac{25}{100} + \frac{1}{100}\left[100 - \frac{1}{2}(400-225)\right] \\
&= \frac{50}{200} + \frac{25}{200} \\
&= \frac{3}{8}.
\end{aligned}$$

(b) Note that Y is non-deterministic, takes values in $[10,20]$ and that the function $x \mapsto 1/x$ is strictly convex on $[10,20]$. Hence we can apply Jensen's inequality to get

$$\mathbb{E}\left[\frac{X}{Y}\right] = \mathbb{E}[X]\,\mathbb{E}\left[\frac{1}{Y}\right] > \mathbb{E}[X]\,\frac{1}{\mathbb{E}[Y]} = 1,$$

where we used in the first equality that X and Y are independent and in the second equality that X and Y have the same distribution. We can not conclude from this inequality that, on average, Peter arrives at the bus stop later than the bus. Since X and Y have the same distribution, we also have $\mathbb{E}[X] = \mathbb{E}[Y]$.

7. (10 points) **Posterior Distribution**

Let X be a normally distributed random variable with mean θ, which is unknown, and variance equal to 1. Assume that we have a prior distribution for θ, which is the standard normal distribution. Moreover, suppose that in an experiment we observe that $X = x$.

(a) (7 points) *Show that the posterior distribution of θ given $X = x$ is the normal distribution with mean $\frac{x}{2}$ and variance $\frac{1}{2}$.*

(b) (3 points) *Determine the Maximum a Posteriori estimate of θ given $X = x$.*

Solution

(a) The density $f_{X\,|\,\theta=\vartheta}$ of $X\,|\,\theta = \vartheta$ is given by

$$f_{X\,|\,\theta=\vartheta}(x\,|\,\vartheta) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(x-\vartheta)^2}.$$

The density f_θ of θ is given by

$$f_\theta(\vartheta) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\vartheta^2}.$$

Hence for the joint density $f_{X,\theta}$ of X and θ we get

$$\begin{aligned} f_{X,\theta}(x,\vartheta) &= f_{X\,|\,\theta=\vartheta}(x\,|\,\vartheta)f_\theta(\vartheta) \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(x-\vartheta)^2-\frac{1}{2}\vartheta^2} \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}2(\vartheta^2-x\vartheta+\frac{1}{2}x^2)} \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}2(\vartheta-\frac{1}{2}x)^2-\frac{1}{4}x^2}. \end{aligned}$$

The marginal density f_X of X is obtained by integrating the joint density $f_{X,\theta}(x,\vartheta)$ with respect to ϑ:

$$f_X(x) = \int_{\mathbb{R}} f_{X,\theta}(x,\vartheta)\,d\vartheta = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{4}x^2}\sqrt{\frac{1}{2}}\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi\frac{1}{2}}}\mathrm{e}^{-\frac{1}{2}2(\vartheta-\frac{1}{2}x)^2}\,d\vartheta = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{4}x^2}\sqrt{\frac{1}{2}}.$$

Conditioned on $X = x$, θ has the density

$$f_{\theta\,|\,X=x}(\vartheta\,|\,x) = \frac{f_{X,\theta}(x,\vartheta)}{f_X(x)} = \frac{1}{\sqrt{2\pi\frac{1}{2}}}\mathrm{e}^{-\frac{1}{2}2(\vartheta-\frac{1}{2}x)^2}.$$

We conclude that

$$\theta\,|\,X = x \sim \mathcal{N}\left(\frac{x}{2},\frac{1}{2}\right).$$

(b) The Maximum a Posteriori estimate $\hat{\theta}_{MAP}$ of θ given $X = x$ is defined as

$$\hat{\theta}_{MAP} := \arg\max_{\vartheta \in \mathbb{R}} f_{\theta \,|\, X=x}(\vartheta \,|\, x) = \arg\max_{\vartheta \in \mathbb{R}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2} 2 (\vartheta - \frac{1}{2} x)^2}.$$

Since $-\frac{1}{2}\, 2\, (\vartheta - \frac{1}{2}\, x)^2 < 0$ for all $\vartheta \in \mathbb{R} \setminus \{\frac{1}{2}\, x\}$ and $-\frac{1}{2}\, 2\, (\frac{1}{2}\, x - \frac{1}{2}\, x)^2 = 0$, we get

$$\hat{\theta}_{MAP} = \frac{x}{2}.$$

8. (10 points) **Maximum Likelihood Estimation**

Let $\theta > -2$ be unknown. Suppose X is a random variable with density f_θ given by

$$f_\theta(x) = \begin{cases} (\theta+2)x^{\theta+1} & \text{if } 0 \le x \le 1, \\ 0 & \text{else.} \end{cases}$$

Let $(X_i)_{i\ge 1}$ be independent, identically distributed random variables with the same distribution as X.

(a) (7 points) Let us consider $X_1, \dots, X_n$. *Show that the Maximum Likelihood Estimator (MLE)* $\hat{\theta}_{MLE}(X_1, \dots, X_n)$ *of* θ *is*

$$\hat{\theta}_{MLE}(X_1, \dots, X_n) = -2 - \frac{1}{\frac{1}{n}\sum_{i=1}^n \log(X_i)}.$$

(b) (3 points) For the random variable X one can show that

$$\mathbb{E}[\log(X)] = -\frac{1}{\theta+2} \qquad \text{and} \qquad \mathrm{Var}[\log(X)] < \infty.$$

Use these two results to show that $\hat{\theta}_{MLE}(X_1, \dots, X_n)$ *given in (a) converges almost surely to* θ*, as* $n \to \infty$.

Solution

(a) The likelihood function $L(\theta)$ is given by

$$L(\theta) = \prod_{i=1}^n f_\theta(X_i) = \prod_{i=1}^n (\theta+2)X_i^{\theta+1} \cdot \mathbf{1}_{\{0\le X_i \le 1\}} = \prod_{i=1}^n (\theta+2)X_i^{\theta+1},$$

since $X_1, \dots, X_n$ only take values in $[0,1]$. By taking the logarithm, we get the log-likelihood function

$$l(\theta) = \log[L(\theta)] = \log\left[\prod_{i=1}^n (\theta+2)X_i^{\theta+1}\right] = \sum_{i=1}^n \left[\log(\theta+2) + (\theta+1)\log(X_i)\right].$$

We then have

$$\hat{\theta}_{MLE}(X_1, \dots, X_n) = \arg\max_{\theta>-2} L(\theta) = \arg\max_{\theta>-2} l(\theta).$$

We find the maximum of $l(\theta)$ on $(-2,\infty)$ by differentiating $l(\theta)$ and setting the derivative equal to zero. We have

$$\begin{aligned}\frac{\partial}{\partial\theta} l(\theta) &= \frac{\partial}{\partial\theta}\sum_{i=1}^n \left[\log(\theta+2) + (\theta+1)\log(X_i)\right] = \sum_{i=1}^n \left[\frac{1}{\theta+2} + \log(X_i)\right] \\ &= \frac{n}{\theta+2} + \sum_{i=1}^n \log(X_i)\end{aligned}$$

and

$$\frac{n}{\theta+2} + \sum_{i=1}^{n} \log(X_i) = 0 \quad \iff \quad \frac{1}{\theta+2} = -\frac{1}{n}\sum_{i=1}^{n} \log(X_i)$$
$$\iff \quad \theta = -2 - \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log(X_i)}.$$

Since

$$\frac{\partial^2}{\partial\theta^2} l(\theta) = \frac{\partial}{\partial\theta}\left[\frac{n}{\theta+2} + \sum_{i=1}^{n} \log(X_i)\right] = -\frac{n}{(\theta+2)^2} < 0 \qquad \text{for all } \theta > -2,$$

$l(\theta)$ is strictly concave on $(-2, \infty)$. Therefore the unique root of the log-likelihood function found above is indeed the location of the maximum of the log-likelihood function. Hence we have found

$$\hat{\theta}_{MLE}(X_1, \ldots, X_n) = -2 - \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log(X_i)}.$$

(b) Since $[\log(X_i)]_{i \geq 1}$ are i.i.d. random variables and $\mathrm{Var}[\log(X)] < \infty$ according to the hint, we can apply the Strong Law of Large Numbers to get

$$\frac{1}{n}\sum_{i=1}^{n} \log(X_i) \longrightarrow \mathbb{E}[\log(X)] = -\frac{1}{\theta+2}$$

almost surely, as $n \to \infty$, and thus

$$\hat{\theta}_{MLE}(X_1, \ldots, X_n) \longrightarrow -2 - \frac{1}{-\frac{1}{\theta+2}} = \theta$$

almost surely, as $n \to \infty$.

9. (10 points) **Hypothesis Test**

Let $X_1, \dots, X_6$ be the annual average temperature of the last six years. Suppose that $X_1, \dots, X_6$ are independent, normally distributed random variables with unknown mean μ and variance $\sigma^2 = \frac{3}{2}$ and that we made the following observations:

$$X_1 = 10, \qquad X_2 = 12, \qquad X_3 = 11, \qquad X_4 = 9, \qquad X_5 = 13, \qquad X_6 = 11.$$

A meteorologist claims that the expected annual average temperature is 10. We will investigate this statement using a z-test with significance level $\alpha = 0.05$. To begin with, we define the null hypothesis and the alternative hypothesis as

$$H_0: \quad \mu = 10, \qquad H_1: \quad \mu \neq 10.$$

(a) (4 points) *Write down a test statistic T that depends on $X_1, \dots, X_6$ and such that $T \sim \mathcal{N}(0,1)$ under H_0.*

(b) (4 points) Let $\mathbb{P}_0$ denote the probability measure under H_0. *Find $q^* > 0$ such that $\mathbb{P}_0(T \in [-q^*, q^*]) = 1 - \alpha$.*
Hint: You may use the table for the standard normal distribution to read off q^* in the end.

(c) (2 points) *Can we reject H_0 on the basis of the six observations* $10, 12, 11, 9, 13, 11$ *using a significance level of $\alpha = 0.05$?*

Solution

(a) We define $\bar{X} := \frac{1}{6}\sum_{i=1}^{6} X_i$. Since $X_1, \dots, X_6 \overset{iid}{\sim} \mathcal{N}\left(10, \frac{3}{2}\right)$ under H_0, we have that $\bar{X}$ is again normally distributed with

$$\mathbb{E}\left[\bar{X}\right] = \frac{1}{6}\sum_{i=1}^{6} \mathbb{E}[X_i] = 10 \quad \text{and} \quad \operatorname{Var}\left(\bar{X}\right) = \frac{1}{36}\sum_{i=1}^{6} \operatorname{Var}(X_i) = \frac{1}{4}$$

under H_0. If we then define

$$T := T(X_1, \dots, X_6) := \frac{\bar{X} - \mathbb{E}\left[\bar{X}\right]}{\sqrt{\operatorname{Var}\left(\bar{X}\right)}} = \frac{\bar{X} - 10}{\frac{1}{2}} = 2\bar{X} - 20,$$

we have that $T \sim \mathcal{N}(0,1)$ under H_0.

(b) Let $q > 0$. Since the normal distribution is symmetric, we can calculate

$$\begin{aligned}\mathbb{P}_0(T \in [-q, q]) &= \mathbb{P}_0(T \leq q) - \mathbb{P}_0(T < -q) = \mathbb{P}_0(T \leq q) - \mathbb{P}_0(T > q) \\ &= \mathbb{P}_0(T \leq q) - (1 - \mathbb{P}_0(T \leq q)) = 2\mathbb{P}_0(T \leq q) - 1.\end{aligned}$$

Therefore we can deduce that

$$\mathbb{P}_0(T \in [-q, q]) = 1 - \alpha \iff 2\mathbb{P}_0(T \leq q) - 1 = 1 - \alpha \iff \mathbb{P}_0(T \leq q) = 1 - \frac{\alpha}{2}.$$

If we write Φ for the distribution function of the standard normal distribution and using that $\alpha = 0.05$, we get

$$q^* = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = \Phi^{-1}(0.975).$$

Checking the table for the standard normal distribution, we find $q^* = 1.96$.

(c) We use the observations given in the exercise to calculate

$$T(10, 12, 11, 9, 13, 11) = 2 \cdot \frac{1}{6} \cdot 66 - 20 = 2.$$

Since $2 > q^* = 1.96$, we reject H_0 on the basis of the six observations $10, 12, 11, 9, 13, 11$ using a significance level of $\alpha = 0.05$.

10. (10 points) **Confidence Interval**

Suppose we have a coin with $\mathbb{P}(\text{"head"}) = p$, where $p \in (0,1)$ is unknown and we toss this coin independently n times. For all $1 \le i \le n$, we define

$$X_i := \begin{cases} 1 & \text{if we get head in the i-th throw,} \\ 0 & \text{if we get tail in the i-th throw.} \end{cases}$$

Moreover, we define $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$. Finally, let $\alpha \in (0,1)$ and Φ be the standard normal distribution function. *Use the Central Limit Theorem to show that*

$$\left[\bar{X} - \frac{\Phi^{-1}\left(1-\frac{\alpha}{2}\right)}{2\sqrt{n}}, \bar{X} + \frac{\Phi^{-1}\left(1-\frac{\alpha}{2}\right)}{2\sqrt{n}}\right]$$

is an approximate $(1-\alpha)$*-confidence interval for* p.

Hint: The confidence interval resulting from the Central Limit Theorem approximation will depend on p*. Since a confidence interval for* p *is not allowed to depend on* p*, you will have to maximize the interval with respect to* p *to get boundaries that do not depend on* p *anymore.*

Solution

We have that $X_1, \dots, X_n \overset{iid}{\sim} \text{Ber}(p)$, hence

$$\mathbb{E}\left[\bar{X}\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = p \quad \text{and} \quad \text{Var}\left(\bar{X}\right) = \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i) = \frac{p(1-p)}{n}.$$

According to the Central Limit Theorem, we have

$$\frac{\bar{X} - \mathbb{E}\left[\bar{X}\right]}{\sqrt{\text{Var}\left(\bar{X}\right)}} = \frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \approx\sim \mathcal{N}(0,1).$$

Let $Z \sim \mathcal{N}(0,1)$ and $q > 0$. Since the normal distribution is symmetric, we can calculate

$$\begin{aligned} \mathbb{P}(Z \in [-q,q]) &= \mathbb{P}(Z \le q) - \mathbb{P}(Z < -q) = \mathbb{P}(Z \le q) - \mathbb{P}(Z > q) \\ &= \mathbb{P}(Z \le q) - (1 - \mathbb{P}(Z \le q)) = 2\mathbb{P}(Z \le q) - 1 \\ &= 2\Phi(q) - 1. \end{aligned}$$

Thus, we have

$$\mathbb{P}\left\{Z \in \left[-\Phi^{-1}\left(1-\frac{\alpha}{2}\right), \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right]\right\} = 2\Phi\left[\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right] - 1 = 1-\alpha.$$

Therefore,

$$\mathbb{P}\left\{\frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \in \left[-\Phi^{-1}\left(1-\frac{\alpha}{2}\right), \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right]\right\} \approx 1-\alpha,$$

which leads to

$$\mathbb{P}\left[\bar{X} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{p(1-p)}{n}} \leq p \leq \bar{X} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{p(1-p)}{n}}\right] \approx 1 - \alpha.$$

Since a confidence interval of p is not allowed to depend on p, we have to maximize $p(1-p) = p - p^2$ with respect to p. We have

$$\frac{\partial}{\partial p}(p - p^2) = 0 \quad \Longleftrightarrow \quad p = \frac{1}{2} \qquad \text{and} \qquad \frac{\partial^2}{\partial p^2}(p - p^2) = -2 < 0.$$

Hence, $p(1-p)$ is a strictly concave function on $[0, 1]$ with a unique maximum of $\frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}$ at $p = \frac{1}{2}$. We conclude that

$$\left[\bar{X} - \frac{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}{2\sqrt{n}}, \bar{X} + \frac{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}{2\sqrt{n}}\right]$$

is an approximate $(1-\alpha)$-confidence interval for p.

Standard normal (cumulative) distribution function.

$$P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \mathrm{d}y, \quad \text{for } x \geqslant 0$$

	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6408	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.998650	.998694	.998736	.998777	.998817	.998856	.998893	.998930	.998965	.998999
3.1	.999032	.999065	.999096	.999126	.999155	.999184	.999211	.999238	.999264	.999289
3.2	.999313	.999336	.999359	.999381	.999402	.999423	.999443	.999462	.999481	.999499
3.3	.999517	.999534	.999550	.999566	.999581	.999596	.999610	.999624	.999638	.999651
3.4	.999663	.999675	.999687	.999698	.999709	.999720	.999730	.999740	.999749	.999758
3.5	.999767	.999776	.999784	.999792	.999800	.999807	.999815	.999822	.999828	.999835
3.6	.999841	.999847	.999853	.999858	.999864	.999869	.999874	.999879	.999883	.999888
3.7	.999892	.999896	.999900	.999904	.999908	.999912	.999915	.999918	.999922	.999925
3.8	.999928	.999931	.999933	.999936	.999938	.999941	.999943	.999946	.999948	.999950
3.9	.999952	.999954	.999956	.999958	.999959	.999961	.999963	.999964	.999966	.999967