Probability and Statistics
FS 2019
Session Exam
16.08.2019
Time Limit: 180 Minutes

Name: ____________________

Student ID: ____________________

This exam contains 25 pages (including this cover page) and 10 problems. A formulae sheet is provided with the exam.

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

Question	Points	Score
1	8	
2	10	
3	8	
4	10	
5	12	
6	10	
7	10	
8	9	
9	11	
10	12	
Total:	100	

1. (8 points) An urn contains 6 distinct balls, of which 2 are red, 2 are white and 2 are black. You take 3 balls out of the urn, at random and without replacement.

   (a) (3 points) Compute the probability that the selected balls have three different colors.

   (b) (3 points) Compute the probability that the selected balls are of exactly 2 different colors.

   (c) (2 points) Compute the probability that none of the selected balls is white.

   Solution

   (a) Let A be the event that the balls have three different colors. For A to happen, one of balls must be red, one black and one white. There are 2 options for the red ball, 2 for the black and 2 for the white, giving 8 possibilities altogether. On the other hand, since there are 6 balls in total, there are $\binom{6}{3} = 20$ subsets of 3 balls that one could draw, and each occurs with the same probability. Therefore, the probability that the selected balls have three different colors is

   $$P(A) = \frac{8}{20} = \frac{2}{5} = 0.4.$$

   (b) Since there are only 2 balls of each color, it is not possible that all three balls have the same color. Therefore, either the balls are of exactly 2 different colors or they are all of different colors. In other words, if B is the event that the balls are of exactly 2 colors, then $B = A^c$ and

   $$P(B) = 1 - P(A) = \frac{3}{5} = 0.6.$$

   (c) Let C be the event that none of the selected balls is white. For C to happen, 3 balls must be chosen from among the 4 that are not white. There are $\binom{4}{3} = 4$ possibilities for such a choice. As before, there are 20 possibilities in total, and therefore

   $$P(C) = \frac{4}{20} = \frac{1}{5} = 0.2.$$

2. (10 points) 3 friends are sitting at a café. Each of them has a coin, with the following properties: for some $p \in (0,1)$,

- The coin of friend 1 shows heads with probability p.
- The coin of friend 2 also shows heads with probability p.
- The coin of friend 3 shows heads with probability $1-p$.

Each of the friends throws his or her coin. We assume that the outcomes of the tosses are independent.

(a) (3 points) What is the probability that one of the friends is the "odd one out", i.e. his/her coin shows a different outcome than the other two?

(b) (3 points) For what value of p is the probability you computed in (a) minimal? Does this result make sense?

(c) (4 points) Take $p = \frac{1}{2}$. The three friends play the following game: the odd one out will pay for the drinks. If nobody is the odd one out, then they will toss their coins again and again until a decision can be made.

What is the probability that they need at least 4 rounds of tosses to know who is going to pay?

Solution

(a) Either the three friends obtain the same result or one of them is the odd one out, since there are only two possible outcomes (heads or tails). Let A be the event that all three coins show the same outcome, and B be the event that one of the friends is the odd one out. Let A_1 be the event that all three coins show heads and A_2 be the event that all three coins show tails. Then $A = A_1 \cup A_2$ is a disjoint union, so that

$$P(A) = P(A_1) + P(A_2) = p^2(1-p) + p(1-p)^2 = p(1-p)(p+1-p) = p(1-p),$$

using the given probabilities and independence of the coins.
Therefore, the desired probability is

$$q = P(B) = 1 - P(A) = 1 - p(1-p).$$

(b) Taking a derivative and setting it to 0, we have

$$\begin{aligned} \frac{d}{dp}[1 - p(1-p)]_{p=\hat{p}} &= 0 \\ \Leftrightarrow 2\hat{p} - 1 &= 0 \\ \Leftrightarrow \hat{p} &= \frac{1}{2}. \end{aligned}$$

Since the second derivative $(= 2)$ is positive, $\frac{1}{2}$ is the minimiser. In that case, the probability is $q = 1 - \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{3}{4}$. This result makes intuitive sense: if p were close to 1, the first two friends would be likely to obtain heads while the third would be likely to obtain tails, making it likely that the third friend would be the odd one out. Similarly, if p were close to 0, the first two friends would be likely to obtain tails while the third would be likely to obtain heads, again making it likely for the third friend to be the odd one out.

(c) If none of the three friends is the odd one out, they repeat the experiment independently from any previous ones, and with the same probability of reaching a conclusion at that stage ($q = \frac{3}{4}$). Therefore, the number of tosses necessary until a conclusion can be reached follows a geometric distribution, $X \sim \text{Geo}(q)$, where q is the probability that a decision is reached at each new round of tosses.

We can then compute explicitly the probability that they will need at least 4 tosses:

$$\begin{aligned}
P(X \geq 4) &= 1 - P(X \leq 3) \\
&= 1 - P(X = 1) - P(X = 2) - P(X = 3) \\
&= 1 - q - q(1 - q) - q(1 - q)^2 \\
&= 1 - \frac{3}{4} - \frac{3}{4} \times \frac{1}{4} - \frac{3}{4} \times \frac{1^2}{4^2} \\
&= 1 - \frac{3}{4} - \frac{3}{16} - \frac{3}{64} \\
&= \frac{1}{64}.
\end{aligned}$$

3. (8 points) Among students who attended a Probability and Statistics course, some were happy and some were not. Moreover, some dropped out of the course before the end.

   For a randomly chosen student S, consider the events:

$$\begin{aligned} L &= \{\text{the student } S \text{ liked the course}\} \\ L^c &= \{\text{the student } S \text{ did not like the course}\} \\ D &= \{\text{the student } S \text{ dropped the course early}\} \\ D^c &= \{\text{the student } S \text{ attended the course until the end of the semester}\}. \end{aligned}$$

   You are given the following information:

$$\begin{aligned} P(D) &= 0.20, \\ P(L \mid D^c) &= 0.90, \\ P(L^c \mid D) &= 0.80. \end{aligned}$$

   (a) (3 points) What is the probability that the student S liked the course?

   (b) (3 points) You meet Maria in the hall and ask her whether she liked the course. Given that the answer is "yes", what is the conditional probability that she attended until the end of the semester?

   (c) (2 points) Simon did not like the course. What is the conditional probability that he dropped the course early?

   Solution

   (a) By partitioning L into $L \cap D$ and $L \cap D^c$ (which are disjoint and whose union is L), and then using the definition of conditional probability, we obtain

$$\begin{aligned} P(L) &= P(L \cap D) + P(L \cap D^c) \\ &= P(L \mid D)P(D) + P(L \mid D^c)P(D^c). \end{aligned}$$

   Now, $P(D) = 0.20$ and $P(L \mid D^c) = 0.90$ are given. Moreover, $P(D^c) = 1 - P(D) = 0.80$ and

$$P(L \mid D) = \frac{P(L \cap D)}{P(D)} = \frac{P(D) - P(L^c \cap D)}{P(D)} = 1 - P(L^c \mid D) = 0.20,$$

   since $L \cap D$ and $L^c \cap D$ partition D.
   Therefore,

$$P(L) = 0.20 \times 0.20 + 0.90 \times 0.80 = 0.76 = \frac{19}{25}.$$

   (b) We are seeking the conditional probability of D^c given L, i.e. $P(D^c \mid L)$. By Bayes' theorem, this is given by

$$P(D^c \mid L) = \frac{P(L \mid D^c)P(D^c)}{P(L)} = \frac{0.90 \times 0.80}{0.76} = \frac{18}{19}.$$

(c) Similarly, we seek $P(D \mid L^c)$. Note that $P(L^c) = 1 - P(L) = 0.24$. By Bayes' theorem, we obtain

$$P(D \mid L^c) = \frac{P(L^c \mid D)P(D)}{P(L^c)} = \frac{0.80 \times 0.20}{0.24} = \frac{2}{3}.$$

4. (10 points) You are browsing mathematics textbooks by two publishers, publisher A and publisher B.

   (a) (2 points) You pick up a book from publisher A. You know that books from publisher A have, on average, 1 typographical error per 50 pages.

   What is an appropriate model for the number of errors in the book you picked up? Specify the parameter of this model.

   (b) (3 points) If this book has 200 pages, give the probability that it contains at least 2 errors.

   (c) (2 points) Give the probability that it contains no errors in its last 40 pages.

   (d) (3 points) Maths books from publisher B contain on average 1 typographical error per 80 pages. Along with your 200 page book from publisher A, you also pick up a 400 page book from publisher B. The numbers of errors in these two books are independent.
   What is the probability that these two books (combined) have strictly fewer than 4 errors?

Solution

(a) An appropriate model for the number of errors is a Poisson distribution, $\text{Poi}(\lambda)$. If the textbook has n pages, it will have on average $\frac{n}{50}$ errors, so $\lambda = \frac{n}{50}$ (as we know that the expectation of the $\text{Poi}(\lambda)$ distribution is λ).

(b) In this case, the number X of errors out of 200 pages has a $\text{Poi}\left(\frac{200}{50}\right) = \text{Poi}(4)$ distribution. Therefore,

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \frac{4^0 e^{-4}}{0!} - \frac{4^1 e^{-4}}{1!} \\ &= 1 - 5e^{-4}. \end{aligned}$$

(c) The expected number of errors is $\frac{40}{50} = 0.8$. Thus, the number of errors $X \sim \text{Poi}(0.8)$ and therefore

$$P(X = 0) = e^{-0.8}.$$

(d) As in part (b), the number of errors in the textbook from publisher A has a $\text{Poi}(4)$ distribution. Furthermore, the expected number of errors in the textbook from publisher B is $\frac{400}{80} = 5$, and so the number of errors $Y \sim \text{Poi}(5)$. Since the numbers of errors in the books are independent, their combined number of errors is $X + Y \sim \text{Poi}(9)$, by standard properties of the Poisson distribution. Thus,

$$\begin{aligned} P(X + Y < 4) &= P(X + Y = 0) + P(X + Y = 1) + P(X + Y = 2) + P(X + Y = 3) \\ &= \frac{9^0 e^{-9}}{0!} + \frac{9^1 e^{-9}}{1!} + \frac{9^2 e^{-9}}{2!} + \frac{9^3 e^{-9}}{3!} \\ &= e^{-9} + 9e^{-9} + \frac{81}{2} e^{-9} + \frac{243}{2} e^{-9} \\ &= 172 e^{-9}. \end{aligned}$$

5. (12 points) A random vector $(X,Y)^T$ takes values in $\mathbb{N}_0 \times \mathbb{N}_0$, and has joint pmf given by

$$p(x,y) = \frac{c}{2^{\max(x,y)}}, \quad (x,y) \in \mathbb{N}_0 \times \mathbb{N}_0$$

for some $c > 0$.

(a) (3 points) Show that

$$\sum_{x\in\mathbb{N}_0, y\in\mathbb{N}_0} \frac{1}{2^{\max(x,y)}} = 2\left(1 + \sum_{y=1}^{\infty} \frac{y}{2^y}\right).$$

Hint. Consider the cases $x \neq y$, $x = y$ separately.

(b) (3 points) Let $Z \sim \mathrm{Geo}(p)$ for some $p \in (0,1)$. Using standard results on convergence and differentiation of power series, show that

$$E(Z) = 1/p.$$

Deduce the value of the normalising constant c, by taking $p = 1/2$.

Remark. You **may not** quote from the formulae sheet for this part!

(c) (3 points) Find the marginal pmf of X and that of Y. Are X and Y independent random variables? Justify your answer.

(d) (3 points) Taking again $Z \sim \mathrm{Geo}(p)$, and again using standard results on convergence and differentiation of power series, show that

$$E[Z(Z-1)] = 2(1-p)/p^2.$$

Deduce E(X) and E(Y) by taking $p = 1/2$.

Remark. You **may not** quote from the formulae sheet for this part!

Solution

(a) We will split the sum by cases $x < y$, $x = y$ and $x > y$. We obtain the following:

$$\begin{aligned}
\sum_{x\in\mathbb{N}_0, y\in\mathbb{N}_0} \frac{1}{2^{\max(x,y)}} &= \sum_{x,y\in\mathbb{N}_0, x>y} \frac{1}{2^{\max(x,y)}} + \sum_{x,y\in\mathbb{N}_0, x=y} \frac{1}{2^{\max(x,y)}} + \sum_{x,y\in\mathbb{N}_0, x<y} \frac{1}{2^{\max(x,y)}} \\
&= \sum_{x,y\in\mathbb{N}_0, x>y} \frac{1}{2^x} + \sum_{x,y\in\mathbb{N}_0, x=y} \frac{1}{2^x} + \sum_{x,y\in\mathbb{N}_0, x<y} \frac{1}{2^y} \\
&= \sum_{x=0}^{\infty} \frac{1}{2^x}\left(\sum_{y=0}^{x-1} 1\right) + \sum_{x=0}^{\infty} \frac{1}{2^x}\left(\sum_{y=x}^{x} 1\right) + \sum_{y=0}^{\infty} \frac{1}{2^y}\left(\sum_{x=0}^{y-1} 1\right) \\
&= \sum_{x=0}^{\infty} \frac{x}{2^x} + \sum_{x=0}^{\infty} \frac{1}{2^x} + \sum_{y=0}^{\infty} \frac{y}{2^y}
\end{aligned}$$

Changing labels, it is clear that $\sum_{x=0}^{\infty} \frac{x}{2^x} = \sum_{y=0}^{\infty} \frac{y}{2^y}$. Moreover, $\sum_{x=0}^{\infty} \frac{1}{2^x} = \frac{1}{1-\frac{1}{2}} = 2$ as it is a geometric series. Thus,

$$\sum_{x\in\mathbb{N}_0, y\in\mathbb{N}_0} \frac{1}{2^{\max(x,y)}} = 2\left(1+\sum_{y=1}^{\infty}\frac{y}{2^y}\right)$$

as we wanted.

(b) We know that the geometric series

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$$

converges absolutely for any $z \in \mathbb{C}$ with $|z| < 1$. By standard results from complex analysis, it follows that this series is analytic on the disk defined by $|z| < 1$, and moreover its derivative is

$$\begin{aligned} \frac{d}{dz}\left(\sum_{k=0}^{\infty} z^k\right) &= \frac{d}{dz}\left(\frac{1}{1-z}\right) \\ \Leftrightarrow \sum_{k=1}^{\infty} kz^{k-1} &= \frac{1}{(1-z)^2}, \end{aligned}$$

for any z with $|z| < 1$.

To compute the expectation of $Z \sim \mathrm{Geo}(p)$, we have

$$\begin{aligned} E[Z] &= \sum_{k=1}^{\infty} kp_Z(k) \\ &= \sum_{k=1}^{\infty} kp(1-p)^{k-1} \\ &= p\sum_{k=1}^{\infty} k(1-p)^{k-1} \\ &= p\frac{1}{(1-(1-p))^2} \\ &= \frac{1}{p}, \end{aligned}$$

since $1-p \in (0,1)$ and by part (a). In particular, if $p = \frac{1}{2}$, then $E(Z) = 2$.

To determine c, note that

$$\begin{aligned} 1 &= \sum_{x\in\mathbb{N}_0, y\in\mathbb{N}_0} p(x,y) \\ &= \sum_{x\in\mathbb{N}_0, y\in\mathbb{N}_0} \frac{c}{2^{\max(x,y)}} \\ &= 2c\left(1+\sum_{y=1}^{\infty}\frac{y}{2^y}\right). \end{aligned}$$

We see that the sum is given by the same expression as the expectation of Z for $p = \frac{1}{2}$, so

$$\begin{aligned} 1 &= 2c(1+2) \\ \Leftrightarrow c &= \frac{1}{6}. \end{aligned}$$

(c) The marginal pmf of X can be determined as follows: for $x \in \mathbb{N}_0$,

$$\begin{aligned} p_X(x) &= \sum_{y \in \mathbb{N}_0} p(x,y) \\ &= \sum_{y \in \mathbb{N}_0} \frac{1}{6 \times 2^{\max(x,y)}} \\ &= \sum_{y \in \mathbb{N}_0, y<x} \frac{1}{6 \times 2^x} + \sum_{y \in \mathbb{N}_0, y \geq x} \frac{1}{6 \times 2^y} \\ &= \frac{x}{6 \times 2^x} + \frac{1}{6 \times 2^x} \sum_{k=0}^{\infty} \frac{1}{2^k} \\ &= \frac{x}{6 \times 2^x} + \frac{2}{6 \times 2^x} \\ &= \frac{x+2}{6 \times 2^x}. \end{aligned}$$

Since the pmf $p(x,y)$ is symmetric in x and y, we immediately deduce by symmetry that

$$p_Y(y) = \frac{y+2}{6 \times 2^y}$$

for $y \in \mathbb{N}_0$.
Clearly, X and Y are not independent, since that would imply that

$$p(x,y) = \frac{1}{6 \times 2^{\max(x,y)}} = p_X(x) p_Y(y) = \frac{x+2}{6 \times 2^x} \frac{y+2}{6 \times 2^y},$$

which is not the case (e.g. $p(0,0) = \frac{1}{6}$ while $p_X(0)p_Y(0) = \frac{1}{9}$).

(d) Using the geometric series again as in part (b), we can take another derivative to obtain

$$\begin{aligned} \frac{d}{dz}\left(\sum_{k=1}^{\infty} k z^{k-1}\right) &= \frac{d}{dz}\left(\frac{1}{(1-z)^2}\right) \\ \Leftrightarrow \left(\sum_{k=2}^{\infty} k(k-1) z^{k-2}\right) &= \frac{2}{(1-z)^3}, \end{aligned}$$

for any z with $|z| < 1$.
For $Z \sim \text{Geo}(p)$, we have

$$
\begin{aligned}
E[Z(Z-1)] &= \sum_{k=1}^{\infty} k(k-1)p_Z(k) \\
&= \sum_{k=2}^{\infty} k(k-1)p(1-p)^{k-1} \\
&= p(1-p)\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} \\
&= p(1-p)\frac{2}{(1-(1-p))^3} \\
&= \frac{2(1-p)}{p^2},
\end{aligned}
$$

since once again $1-p \in (0,1)$. In particular, if $p = \frac{1}{2}$,

$$
E[Z(Z-1)] = \sum_{k=2}^{\infty} k(k-1)\frac{1}{2^k} = 4.
$$

To compute the expectation of X, note that

$$
\begin{aligned}
E[X] &= \sum_{x=1}^{\infty} x p_X(x) \\
&= \sum_{x=1}^{\infty} x \frac{x+2}{6 \times 2^x} \\
&= \frac{1}{6}\sum_{x=2}^{\infty} \frac{x(x-1)}{2^x} + \frac{1}{6}\sum_{x=1}^{\infty} \frac{3x}{2^x} \\
&= \frac{1}{6} \times 4 + \frac{1}{6} \times 3 \times 2 \\
&= \frac{5}{3},
\end{aligned}
$$

using both results from both part (b) and (d) to simplify the sums. By symmetry, $E[Y] = E[X] = \frac{5}{3}$.

6. (10 points) Let X and Y be independent random variables such that $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$, for some $\lambda > 0$ and $\mu > 0$.

   (a) (1 point) Write down the joint density of the random vector $(X, Y)^T$ with respect to Lebesgue measure on $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$.

   (b) (3 points) Find $E(\max(X, Y))$.

   (c) (1 point) Find a simple expression for $\max(X, Y) + \min(X, Y)$ and deduce $E(\min(X, Y))$.

   (d) (3 points) Assuming that $\mu = \lambda = 1$, calculate $P(Y < X + 1)$ and deduce $P(|X - Y| < 1)$.

   (e) (2 points) Taking general $\lambda, \mu > 0$ again, let $Z = X + Y$. Using well-known properties of conditional expectation, find $E(Z \mid X)$ (specifying what properties you use).

Solution

(a) Since X and Y are independent, their joint density $f_{X,Y}$ with respect to the Lebesgue measure is given by the product of the individual densities:

$$f_{X,Y}(x, y) = \lambda\mu e^{-\lambda x} e^{-\mu y} \mathbb{1}_{x,y \geq 0}.$$

(b) We compute this expectation directly, by taking an integral. Since the integrands are non-negative, we freely use Fubini's theorem to exchange integrals as required.

$$\begin{aligned}
E[\max(X, Y)] &= \iint_{\mathbb{R}^2} \max(x, y) f_{X,Y}(x, y)\, dx\, dy \\
&= \iint_{\mathbb{R}^2} \max(x, y) \lambda\mu e^{-\lambda x} e^{-\mu y} \mathbb{1}_{x,y \geq 0}\, dx\, dy \\
&= \int_0^\infty \int_0^\infty \lambda\mu \max(x, y) e^{-\lambda x} e^{-\mu y}\, dx\, dy \\
&= \int_0^\infty \int_0^\infty \lambda\mu e^{-\lambda x} e^{-\mu y} (x \mathbb{1}_{x \geq y} + y \mathbb{1}_{y > x})\, dx\, dy \\
&= \int_0^\infty \int_0^x \lambda\mu x e^{-\lambda x} e^{-\mu y}\, dx\, dy + \int_0^\infty \int_0^y \lambda\mu y e^{-\lambda x} e^{-\mu y}\, dy\, dx \\
&= \int_0^\infty \lambda x e^{-\lambda x} (1 - e^{-\mu x})\, dx + \int_0^\infty \mu y e^{-\mu y} (1 - e^{-\lambda y})\, dy \\
&= \int_0^\infty x \lambda e^{-\lambda x}\, dx - \frac{\lambda}{\lambda + \mu} \int_0^\infty x(\lambda + \mu) e^{-(\lambda+\mu)x}\, dx \\
&\quad + \int_0^\infty y \mu e^{-\mu y}\, dy - \frac{\mu}{\lambda + \mu} \int_0^\infty y(\lambda + \mu) e^{-(\lambda+\mu)y}\, dy \\
&= \frac{1}{\lambda} - \frac{\lambda}{\lambda + \mu} \times \frac{1}{\lambda + \mu} + \frac{1}{\mu} - \frac{\mu}{\lambda + \mu} \times \frac{1}{\lambda + \mu} \\
&= \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu},
\end{aligned}$$

using the formulas for the expectations of random variables with distributions $\text{Exp}(\lambda), \text{Exp}(\mu)$ and $\text{Exp}(\mu + \lambda)$.

(c) Since (for any $\omega \in \Omega$) one of X, Y is the minimum and the other is the maximum, it follows that

$$\max(X, Y) + \min(X, Y) = X + Y.$$

Therefore,

$$\begin{aligned} E(\min(X, Y)) &= E(X) + E(Y) - E(\max(X, Y)) \\ &= \frac{1}{\lambda} + \frac{1}{\mu} - \left(\frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda+\mu}\right) \\ &= \frac{1}{\lambda+\mu}. \end{aligned}$$

(d) Once again, we compute this probability directly through an integral. Let $D = \{x, y \in \mathbb{R}^2 : y < x + 1\}$. As before, we will occasionally use Fubini's theorem. We obtain

$$\begin{aligned} P(Y < X + 1) &= \iint_D f_{X,Y}(x, y)\, dx\, dy \\ &= \int_0^\infty \int_0^\infty \lambda\mu e^{-\lambda x} e^{-\mu y} \mathbb{1}_{y<x+1}\, dx\, dy \\ &= \int_0^\infty \int_0^{x+1} \lambda\mu e^{-\lambda x} e^{-\mu y}\, dx\, dy \\ &= \int_0^\infty \lambda e^{-\lambda x}(1 - e^{-\mu(x+1)})\, dx \\ &= \int_0^\infty \lambda e^{-\lambda x}\, dx - \frac{\lambda e^{-\mu}}{\lambda+\mu} \int_0^\infty (\lambda+\mu) e^{-(\lambda+\mu)x}\, dx \\ &= 1 - \frac{\lambda e^{-\mu}}{\lambda+\mu}, \end{aligned}$$

since the densities of $\mathrm{Exp}(\lambda)$ and $\mathrm{Exp}(\lambda+\mu)$-distributed random variables integrate to 1. In particular, if $\lambda = \mu = 1$,

$$P(Y < X + 1) = 1 - \frac{1}{2e}.$$

To find $P(|X - Y| < 1)$, note that if $\mu = \lambda = 1$, then the distribution of (X, Y) is symmetric. In particular,

$$P(Y < X + 1) = P(X < Y + 1).$$

Moreover, since $\{Y \geq X + 1\}$ together with $\{X \geq Y + 1\}$ form a partition of $\{|X - Y| \geq 1\}$, we obtain

$$\begin{aligned} P(|X - Y| < 1) &= 1 - P(|X - Y| \geq 1) \\ &= 1 - [P(Y \geq X + 1) + P(X \geq Y + 1)] \\ &= 1 - [1 - P(Y < X + 1) + 1 - P(X < Y + 1)] \\ &= 1 - \left(\frac{1}{2e} + \frac{1}{2e}\right) \\ &= 1 - \frac{1}{e}. \end{aligned}$$

(e) Since conditional expectations are additive,

$$E[Z \mid X] = E[X + Y \mid X] = E[X \mid X] + E[Y \mid X].$$

Since X is of course completely determined by itself, $E[X \mid X] = X$. Moreover, since Y and X are independent, $E[Y \mid X] = E[Y] = \frac{1}{\mu}$. Therefore, we conclude that

$$E[Z \mid X] = X + \frac{1}{\mu}.$$

7. (10 points) Suppose that the heights (in cm) of individuals of some population are distributed according to a normal distribution. You are given the following:

   - The average height is 170cm.
   - The proportion of individuals who are taller than 190cm is 5%.

   In the following, Φ denotes the cdf of $\mathcal{N}(0,1)$, and $z_\alpha = \Phi^{-1}(\alpha)$ its α-quantile for $\alpha \in (0,1)$.

   (a) (3 points) Find the standard deviation, σ, of the distribution of the heights.

   (b) (2 points) Let p be the probability that a random individual is taller than 160cm. Write p in terms of some value of Φ which you must determine.

   (c) (2 points) 100 individuals are selected at random. Find an expression for the probability that at least half of them are taller than 160cm.

   (d) (3 points) Assuming that 100 is a large enough sample size, give an approximation for the probability in (c). Justify your answer.

   Solution

   (a) We know that the height in cm, H, has a normal distribution $H \sim \mathcal{N}(170, \sigma^2)$, since we are given the average height. The second condition will allow us to determine the standard deviation. Note that

$$\begin{aligned} 0.05 &= P(H > 190) \\ &= P\left(\frac{H-170}{\sigma} > \frac{20}{\sigma}\right) \\ &= 1 - P\left(\frac{H-170}{\sigma} \leq \frac{20}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{20}{\sigma}\right), \end{aligned}$$

   since $\frac{H-170}{\sigma} \sim \mathcal{N}(0,1)$. Therefore, we can solve

$$\begin{aligned} & \Phi\left(\frac{20}{\sigma}\right) = 0.95 \\ \Rightarrow \quad & \frac{20}{\sigma} = z_{0.95} = 1.65 \\ \Rightarrow \quad & \sigma = \frac{20}{1.65} \approx 12.1. \end{aligned}$$

   (b) The probability is

$$
\begin{aligned}
P(H > 160) &= P\left(\frac{H-170}{\sigma} > \frac{-10}{\sigma}\right) \\
&= 1 - P\left(\frac{H-170}{\sigma} \leq \frac{-10}{\sigma}\right) \\
&= 1 - \Phi\left(\frac{-10}{\sigma}\right) \\
&= 1 - \Phi\left(\frac{-1.65}{2}\right) \\
&= \Phi\left(\frac{1.65}{2}\right) \\
&= \Phi\left(0.825\right) \\
&\approx 0.80.
\end{aligned}
$$

(c) Let $H_1, ..., H_{100}$ be the heights of the individuals, which we assume to be i.i.d. $\sim \mathcal{N}(170, \sigma^2)$.
Let $B_i = \mathbb{1}_{H_i > 160}$. Then, since the H_i are independent, so are the B_i. Moreover, each B_i is a Bernoulli random variable with

$$P(B_i = 1) = P(H_i > 160) = p \approx 0.80.$$

Then, at least 50 individuals are taller than 160cm if and only if

$$X = \sum_{i=1}^{100} B_i \geq 50.$$

Note then that $X \sim \text{Bin}(100, p)$, and therefore

$$
\begin{aligned}
P(X \geq 50) &= \sum_{k=50}^{100} P(X = k) \\
&= \sum_{k=50}^{100} \binom{n}{k} p^k (1-p)^{1-k}.
\end{aligned}
$$

(d) If we assume that 100 is a large enough sample, we can obtain a normal approximation, thanks to the central limit theorem. Note that the B_i are i.i.d. $\sim \text{Ber}(p)$, so that $E(B_i) = p$ and $\text{var}(B_i) = p(1-p)$. Therefore, by the central limit theorem,

$$\frac{\frac{1}{100}\sum_{i=1}^{100} B_i - p}{\sqrt{p(1-p)/100}} \stackrel{\text{d}}{\approx} \mathcal{N}(0,1).$$

Based on this approximation, we obtain

$$\begin{aligned}
P\left(X \geq 50\right) &= P\left(\frac{1}{100}\sum_{i=1}^{100} B_i \geq \frac{1}{2}\right) \\
&= P\left(\frac{1}{100}\sum_{i=1}^{100} B_i - p \geq \frac{1}{2} - p\right) \\
&= P\left(\frac{\frac{1}{100}\sum_{i=1}^{100} B_i - p}{\sqrt{p(1-p)/100}} \geq \frac{\frac{1}{2} - p}{\sqrt{p(1-p)/100}}\right) \\
&\approx 1 - \Phi\left(\frac{\frac{1}{2} - p}{\sqrt{p(1-p)/100}}\right) \\
&= \Phi\left(\frac{p - \frac{1}{2}}{\sqrt{p(1-p)/100}}\right) \\
&\approx \Phi\left(7.5\right).
\end{aligned}$$

8. (9 points) Consider a Gaussian vector $(X_1, X_2, X_3)^T$ with expectation $\mu = (0, 1, -1)^T$ and covariance matrix

$$\Sigma = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 7 & -5 \\ 0 & -5 & 9 \end{pmatrix}.$$

(a) (2 points) Give the marginal densities of X_1 and X_2 with respect to Lebesgue measure on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$.

(b) (2 points) Are X_1 and X_2 independent? Are X_2 and X_3 independent?

(c) (2 points) What is the distribution of $X_1 + X_2$? Give its density with respect to Lebesgue measure on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$.

(d) (3 points) Find all the possible values of $\alpha_0 \in \mathbb{R}$ such that X_2 and $X_2 + \alpha_0 X_3$ satisfy

$$\operatorname{cov}(X_2, X_2 + \alpha_0 X_3) = 0.$$

What can you conclude about X_2 and $X_2 + \alpha_0 X_3$?

Solution

(a) Since $(X_1, X_2, X_3)^T$ is a Gaussian vector, linear combinations of its components (and thus the components themselves) are Gaussian. The expectation vector and covariance matrix give us the respective parameters. Thus we can read off

$$X_1 \sim \mathcal{N}(0, 9) \text{ and } X_2 \sim \mathcal{N}(1, 7).$$

Therefore, the respective densities are

$$f_{X_1}(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{x^2}{18}},$$
$$f_{X_2}(x) = \frac{1}{\sqrt{14\pi}} e^{-\frac{(x-1)^2}{14}}.$$

(b) We know from the exercise sheets that components of a Gaussian vector are independent if and only if their covariance is 0. Once again, we can read the covariances from the covariance matrix:

$$\operatorname{cov}(X_1, X_2) = \Sigma_{12} = 0,$$
$$\operatorname{cov}(X_2, X_3) = \Sigma_{23} = -5.$$

Thus, X_1 and X_2 are independent, but X_2 and X_3 are not.

(c) Since X_1, X_2 are independent, we know from lectures that their sum is a Gaussian, whose expectation is the sum of those of X_1 and X_2, and whose variance is the sum of those of X_1 and X_2. Thus,

$$X_1 + X_2 \sim \mathcal{N}(0 + 1, 9 + 7) = \mathcal{N}(1, 16).$$

The density is

$$f_{X_1+X_2}(x) = \frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-1)^2}{32}}.$$

(d) We can compute

$$\begin{aligned}\mathrm{cov}(X_2, X_2 + \alpha_0 X_3) &= \mathrm{var}(X_2) + \alpha_0 \mathrm{cov}(X_2, X_3)\\ &= 7 - 5\alpha_0.\end{aligned}$$

Thus, $\mathrm{cov}(X_2, X_2 + \alpha_0 X_3) = 0$ if and only if $\alpha_0 = \frac{7}{5}$. Once again, as linear combinations of the components of $(X_1, X_2, X_3)^T$, this means that X_2 and $X_2 + \frac{7}{5}X_3$ are independent.

9. (11 points) Let $X_1, ..., X_n$ be i.i.d $\sim \text{Geo}(\theta_0)$ for some unknown $\theta_0 \in \Theta = (0, 1)$.

   (a) (2 points) Find a moment estimator, $\tilde{\theta}_n$, for θ_0.

   (b) (3 points) Justifying your answer, find the MLE, $\hat{\theta}_n$, for θ_0.

   (c) (3 points) Assuming that sufficient regularity conditions hold for this model, compute the Cramér-Rao Lower Bound.

   (d) (1 point) Again assuming sufficient regularity conditions, explicitly give the result on asymptotic normality for $\hat{\theta}_n$.

   (e) (2 points) Deduce from (d) an asymptotic confidence interval for θ_0 of level 0.95.

Solution

(a) We know that for $X_1 \sim \text{Geo}(\theta_0)$, its expectation is $E[X_1] = \frac{1}{\theta_0}$. To obtain a moment estimator, we approximate the expectation by the empirical mean (which is justified by the strong law of large numbers, since the X_i are i.i.d):

$$\frac{1}{n}\sum_{i=1}^{n} X_i \approx E[X_1] = \frac{1}{\theta_0}$$
$$\Rightarrow \theta_0 \approx \tilde{\theta}_n := \frac{n}{\sum_{i=1}^{n} X_i}.$$

(b) To find the MLE, we need to maximise the likelihood function (or equivalently, the log-likelihood function). By independence:

$$L(\theta) = \prod_{i=1}^{n} \theta(1-\theta)^{X_i - 1},$$

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

To maximise, we set the derivative to 0:

$$\begin{aligned}
& l'(\hat{\theta}_n) = 0 \\
\Leftrightarrow \quad & \frac{n}{\hat{\theta}_n} - \frac{\sum_{i=1}^{n} X_i - n}{1-\hat{\theta}_n} = 0 \\
\Leftrightarrow \quad & n(1-\hat{\theta}_n) - \left(\sum_{i=1}^{n} X_i - n\right)\hat{\theta}_n = 0 \\
\Leftrightarrow \quad & n - \hat{\theta}_n \sum_{i=1}^{n} X_i = 0 \\
\Leftrightarrow \quad & \hat{\theta}_n = \frac{n}{\sum_{i=1}^{n} X_i}.
\end{aligned}$$

In this case we obtain the same as the moment estimator. We check that this is the maximum by taking the second derivative:

$$l''(\theta) = -\frac{n}{\theta^2} - \frac{\sum_{i=1}^n X_i - n}{(1-\theta)^2} < 0$$

since each $X_i \geq 1$, so l is strictly concave and $\hat{\theta}_n$ is the MLE.

(c) To compute the Cramér-Rao lower bound, we need to compute the Fisher information:

$$\begin{aligned}
I(\theta_0) &= E_{\theta_0}\left[-\frac{d^2}{d\theta^2}\log(f_\theta(X_1))\,|_{\theta=\theta_0}\right] \\
&= E_{\theta_0}\left[\frac{1}{\theta_0^2} + \frac{X_1 - 1}{(1-\theta_0)^2}\right] \\
&= \frac{1}{\theta_0^2} + \frac{\frac{1}{\theta_0} - 1}{(1-\theta_0)^2} \\
&= \frac{1}{\theta_0^2} + \frac{1}{\theta_0(1-\theta_0)} \\
&= \frac{1-\theta_0+\theta_0}{\theta_0^2(1-\theta_0)} \\
&= \frac{1}{\theta_0^2(1-\theta_0)}.
\end{aligned}$$

The Cramér-Rao lower bound is then

$$\frac{1}{nI(\theta_0)} = \frac{\theta_0^2(1-\theta_0)}{n}.$$

(d) The asymptotic normality result tells us that

$$\sqrt{nI(\theta_0)}(\hat{\theta}_n - \theta_0) = \frac{\sqrt{n}}{\theta_0\sqrt{1-\theta_0}}(\hat{\theta}_n - \theta_0) \xrightarrow{\text{d}} \mathcal{N}(0,1)$$

as $n \to \infty$.

(e) To find a confidence interval of level 0.95, note that by asymptotic normality,

$$P\left(-z_{1-\frac{0.05}{2}} \leq \frac{\sqrt{n}}{\theta_0\sqrt{1-\theta_0}}(\hat{\theta}_n - \theta_0) \leq z_{1-\frac{0.05}{2}}\right) \approx 0.95.$$

This can be rewritten as

$$P\left(\hat{\theta}_n - \frac{\theta_0\sqrt{1-\theta_0}z_{1-\frac{0.05}{2}}}{\sqrt{n}} \leq \theta_0 \leq \hat{\theta}_n + \frac{\theta_0\sqrt{1-\theta_0}z_{1-\frac{0.05}{2}}}{\sqrt{n}}\right) \approx 0.95.$$

If we assume that n is large enough that $\hat{\theta}_n$ is a good approximation for θ_0, we can replace θ_0 with $\hat{\theta}_n$ in the limits of the interval. Moreover, we know that $z_{0.975} \approx 1.96$.This gives the approximate 0.95-confidence interval:

$$\left[\hat{\theta}_n - \frac{1.96\,\hat{\theta}_n\sqrt{1-\hat{\theta}_n}}{\sqrt{n}}, \hat{\theta}_n + \frac{1.96\,\hat{\theta}_n\sqrt{1-\hat{\theta}_n}}{\sqrt{n}}\right].$$

10. (12 points) Let $X_1, ..., X_n$ be i.i.d. $\sim \mathcal{N}(\theta, 1)$ for some $\theta \in \mathbb{R}$. We want to test

$$H_0 : \theta = 1 \quad \text{versus} \quad H_1 : \theta = \theta_1,$$

for some fixed $\theta_1 > 1$.

Fix $\alpha \in (0, 1)$. We denote once again by Φ the cdf of $\mathcal{N}(0, 1)$ and $z_\gamma = \Phi^{-1}(\gamma)$ for $\gamma \in (0, 1)$.

(a) (3 points) Give the NP-test of level α for this hypothesis testing problem.

(b) (2 points) Give the expression for the power of this test. Show that the power converges to 1 as $n \to +\infty$.

Now, let us exchange the roles of the null and alternative hypotheses. That is, we want to test

$$H_0 : \theta = \theta_1 \quad \text{versus} \quad H_1 : \theta = 1.$$

We still assume that $\theta_1 > 1$.

(c) (3 points) Find the NP-test of level α for this new problem as well as its power.

(d) (2 points) Show that for any $x \in \mathbb{R}$, $\Phi(-x) = 1 - \Phi(x)$. Conclude that $z_\gamma = -z_{1-\gamma}$ for $\gamma \in (0, 1)$.

Remark. You **may not** quote from the lectures for this part!

(e) (2 points) What do you conclude about the powers of the tests obtained in (b) and (c)?

Solution

(a) If L_0 is the likelihood function under H_0 and L_1 is the likelihood function under H_1, the NP-test of level α for this problem has the form

$$\chi_\alpha(\mathbf{x}) = \begin{cases} 1, & \frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} > t_\alpha, \\ \gamma, & \frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} = t_\alpha, \\ 0, & \frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} < t_\alpha, \end{cases}$$

for some threshold $t_\alpha \in \mathbb{R}$.

The inequality can be simplified as follows: by independence, one obtains

$$\begin{aligned}
&\frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} > t_\alpha \\
\Leftrightarrow\quad & \frac{\frac{1}{\sqrt{2\pi}} e^{-\sum_{i=1}^n \frac{(x_i-\theta_1)^2}{2}}}{\frac{1}{\sqrt{2\pi}} e^{-\sum_{i=1}^n \frac{(x_i-1)^2}{2}}} > t_\alpha \\
\Leftrightarrow\quad & e^{\frac{-\sum_{i=1}^n (x_i-\theta_1)^2 + \sum_{i=1}^n (x_i-1)^2}{2}} > t_\alpha \\
\Leftrightarrow\quad & -\sum_{i=1}^n (x_i-\theta_1)^2 + \sum_{i=1}^n (x_i-1)^2 > 2\log(t_\alpha) \\
\Leftrightarrow\quad & -\sum_{i=1}^n (x_i^2 - 2x_i\theta_1 + \theta_1^2) + \sum_{i=1}^n (x_i^2 - 2x_i + 1) > 2\log(t_\alpha) \\
\Leftrightarrow\quad & 2(\theta_1 - 1)\sum_{i=1}^n x_i > 2\log(t_\alpha) + n(\theta_1^2 - 1) \\
\Leftrightarrow\quad & \sum_{i=1}^n x_i > \frac{2\log(t_\alpha) + n(\theta_1^2-1)}{2(\theta_1 - 1)},
\end{aligned}$$

since $\theta_1 > 1$. Therefore, the NP-test can be written in the form

$$\chi_\alpha(\mathbf{x}) = \begin{cases} 1, & \sum_{i=1}^n x_i > k_\alpha, \\ \gamma, & \sum_{i=1}^n x_i = k_\alpha, \\ 0, & \sum_{i=1}^n x_i < k_\alpha, \end{cases}$$

for some $k_\alpha \in \mathbb{R}, \gamma \in [0,1]$. Since, under H_0 or H_1, the random vector $(X_1, ..., X_n)$ has a density with respect to Lebesgue measure on $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, the probability that $\sum_{i=1}^n x_i = k_\alpha$ is 0 (under both H_0 and H_1). Therefore, we may take $\gamma = 0$, say.

In order to ensure that the test has level α, we need to determine k_α by calculating the probability of type I error:

$$\begin{aligned}
\alpha &= E_{H_0}[\chi_\alpha(\mathbf{X})] \\
&= P_{H_0}\left(\sum_{i=1}^n X_i > k_\alpha\right).
\end{aligned}$$

Under H_0, each $X_i \sim \mathcal{N}(1,1)$, so, by independence, $\sum_{i=1}^n X_i \sim \mathcal{N}(n,n)$. Thus,

$$\begin{aligned}
\alpha &= P_{H_0}\left(\frac{\sum_{i=1}^n X_i - n}{\sqrt{n}} > \frac{k_\alpha - n}{\sqrt{n}}\right) \\
&= 1 - \Phi\left(\frac{k_\alpha - n}{\sqrt{n}}\right) \\
\Leftrightarrow z_{1-\alpha} &= \frac{k_\alpha - n}{\sqrt{n}} \\
\Leftrightarrow k_\alpha &= n + \sqrt{n} z_{1-\alpha}.
\end{aligned}$$

For this choice of k_α, the corresponding χ_α is then the NP-test of level α.

(b) To compute the power, we now consider the probability of rejection under H_1. Under H_1, by independence, $\sum_{i=1}^n X_i \sim \mathcal{N}(n\theta_1, n)$. Denoting by β_n the power of the test with a sample size of n,

$$\begin{aligned}\beta_n &= P_{H_1}\left(\sum_{i=1}^n X_i > k_\alpha\right)\\ &= P_{H_1}\left(\frac{\sum_{i=1}^n X_i - n\theta_1}{\sqrt{n}} > \frac{k_\alpha - n\theta_1}{\sqrt{n}}\right)\\ &= 1 - \Phi\left(\frac{k_\alpha - n\theta_1}{\sqrt{n}}\right)\\ &= 1 - \Phi\left(\sqrt{n} - \sqrt{n}\theta_1 + z_{1-\alpha}\right).\end{aligned}$$

As $n \to +\infty$, $\sqrt{n} - \sqrt{n}\theta_1 \to -\infty$ since $\theta_1 > 1$. Therefore, $\Phi\left(\sqrt{n} - \sqrt{n}\theta_1 + z_{1-\alpha}\right) \to 0$ as $n \to \infty$, since for any cdf F we have $\lim_{z\to-\infty} F(z) = 0$. Finally, we conclude that $\beta_n = 1 - \Phi\left(\sqrt{n} - \sqrt{n}\theta_1 + z_{1-\alpha}\right) \to 1$ as $n \to +\infty$.

(c) One can switch θ_1 and 1 in the calculations for part (a), which shows that the NP-test must have the same form, with the exception that, since now $\theta_1 > 1$, the direction of the inequalities is flipped. In other words, the NP-test of level α for this new problem has the form

$$\tilde{\chi}_\alpha(\mathbf{x}) = \begin{cases} 1, & \sum_{i=1}^n x_i < \tilde{k}_\alpha \\ 0, & \sum_{i=1}^n x_i \geq \tilde{k}_\alpha, \end{cases}$$

for some $\tilde{k}_\alpha \in \mathbb{R}$. To determine $\tilde{k}_\alpha$, we find the probability of type I errors. This time, $\sum_{i=1}^n X_i \sim \mathcal{N}(n\theta_1, n)$ under H_0.

$$\begin{aligned}\alpha &= E_{H_0}[\tilde{\chi}_\alpha(\mathbf{X})]\\ &= P_{H_0}\left(\sum_{i=1}^n X_i < \tilde{k}_\alpha\right)\\ &= P_{H_0}\left(\frac{\sum_{i=1}^n X_i - n\theta_1}{\sqrt{n}} < \frac{\tilde{k}_\alpha - n\theta_1}{\sqrt{n}}\right)\\ &= \Phi\left(\frac{\tilde{k}_\alpha - n\theta_1}{\sqrt{n}}\right)\\ \Leftrightarrow z_\alpha &= \frac{\tilde{k}_\alpha - n\theta_1}{\sqrt{n}}\\ \Leftrightarrow \tilde{k}_\alpha &= \sqrt{n}z_\alpha + n\theta_1.\end{aligned}$$

Finally, the power is given by

$$\begin{aligned}\beta &= P_{H_1}\left(\sum_{i=1}^{n} X_i < \tilde{k}_\alpha\right)\\ &= P_{H_1}\left(\frac{\sum_{i=1}^{n} X_i - n}{\sqrt{n}} < \frac{\tilde{k}_\alpha - n}{\sqrt{n}}\right)\\ &= \Phi\left(\frac{\tilde{k}_\alpha - n}{\sqrt{n}}\right)\\ &= \Phi\left(\sqrt{n}\theta_1 - \sqrt{n} + z_\alpha\right).\end{aligned}$$

(d) Let $Z \sim \mathcal{N}(0,1)$. For any $x \in \mathbb{R}$, and using the symmetry of the density of Z (since $e^{-\frac{x^2}{2}} = e^{-\frac{(-x)^2}{2}}$), we obtain

$$\begin{aligned}\Phi(-x) &= P(Z \leq -x)\\ &= \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy\\ &= \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y'^2}{2}}\, dy' \quad (y' = -y)\\ &= P(Z > x)\\ &= 1 - P(Z \leq x)\\ &= 1 - \Phi(x)\end{aligned}$$

(we also used the fact that $P(Z = x) = 0$).
For the second equality, note that

$$\Phi(-z_{1-\gamma}) = 1 - \Phi(z_{1-\gamma}) = 1 - (1-\gamma) = \gamma = \Phi(z_\gamma),$$

and since Φ is strictly increasing, thus injective, we deduce that $-z_{1-\gamma} = z_\gamma$.

(e) Recall that, in part (a), we obtained a power of

$$1 - \Phi\left(\sqrt{n} - \sqrt{n}\theta_1 + z_{1-\alpha}\right)$$

while in part (c) we obtained

$$\Phi\left(\sqrt{n}\theta_1 - \sqrt{n} + z_\alpha\right).$$

We show that these two are the same (assuming that the choice of α and n is the same), using the previous properties. Indeed,

$$\begin{aligned}1 - \Phi\left(\sqrt{n} - \sqrt{n}\theta_1 + z_{1-\alpha}\right) &= \Phi\left(-\sqrt{n} + \sqrt{n}\theta_1 - z_{1-\alpha}\right)\\ &= \Phi\left(-\sqrt{n} + \sqrt{n}\theta_1 + z_\alpha\right),\end{aligned}$$

by the first and then the second properties proven in (d). Thus, the two tests have the same power.

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆