Midterm Probability Theory and Statistics (CSE1210)

May 25th 2023, 9:00 — 11:00

TU Delft, Delft Institute of Applied Mathematics

You can use a non-graphic, non-programmable, non-symbolic calculator. You can use your own formula sheet (1 A4 side).
Questions 11 and 12 are open, so only a final answer will not be graded there. Show your workings.

Multiple choice questions

1. Consider two events E and G. If $P(G|E) = 0.6$, $P(E) = 0.35$ and $P(E \cup G) = 0.8$, then what is $P(G)$?

   (A) 0.46 **(B)** 0.50 **(C)** 0.54 **(D)** 0.58 **(E)** 0.62 **(F)** 0.66

2. A computer program has been developed that tries to recognize animal species from an image. So you upload an image of an animal (so far it only works on mammals, fish and birds) and the program tries to recognize the species, i.e. the output is something like "this is a clownfish".

   If the image shows a mammal, the program has a probability of 80% of finding the correct species. If the image shows a fish, the probability is 50%. If the image shows a bird, the probability is 60%.

   Now I pick an image at random from a database containing 2000 images: 1200 images of mammals, 600 images of fish and 200 images of birds. The program unfortunately fails to correctly identify the species. What is the probability that this image shows a mammal (rounded to two decimal digits)?

   (A) 0.15 **(B)** 0.23 **(C)** 0.31 **(D)** 0.39 **(E)** 0.47 **(F)** 0.55

3. Consider the continuous random variable X which has a beta distribution. We didn't cover this distribution, but you can use that X only takes values on $[0, 1]$, that $E[X] = \frac{1}{3}$ and $\text{Var}(X) = \frac{1}{18}$.

   Recall that by Chebyshev's inequality, for any random variable Y and for any $a > 0$,

   $$P\left(|Y - E[Y]| \geq a\right) \leq \frac{\text{Var}(Y)}{a^2}.$$

   Use Chebyshev's inequality to find the best possible upper bound for $P(X \geq \frac{5}{6})$, that is, find the number that belongs on the dots in $P(X \geq \frac{5}{6}) \leq \ldots$.

   (A) $\frac{4}{18}$ **(B)** $\frac{5}{18}$ **(C)** $\frac{6}{18}$ **(D)** $\frac{7}{18}$ **(E)** $\frac{8}{18}$ **(F)** $\frac{9}{18}$ **(G)** $\frac{10}{18}$

4. Consider 100 independent random variables, each having a beta distribution as in the previous question: $E[X_i] = \frac{1}{3}$ and $\mathrm{Var}(X_i) = \frac{1}{18}$. Approximate the probability that their sum is at least 36.

   (A) 0.1292 **(B)** 0.0869 **(C)** 0.0485 **(D)** 0.0073 **(E)** 0.0011

5. Compute $E[X]$ where X and Y are continuous random variables, where both X and Y only take values in $[1,2]$ and there their joint probability density function is given by $f_{X,Y}(x,y) = \frac{3}{14}(x^2+y^2)$. Round to two decimal digits.

   (A) 1.17 **(B)** 1.28 **(C)** 1.35 **(D)** 1.42 **(E)** 1.55 **(F)** 1.71

Short answer questions

6. Consider a random variable X, determined in the following way. I toss a fair coin. If it shows Heads, then I take $X = 2$. If it shows Tails, then I pick a number from a $U(0,1)$ distribution (the continuous uniform distribution on $[0,1]$) and I let X be this number. Draw the distribution function F_X. Make sure you indicate all relevant numbers on both axes.

7. Consider 10 radioactive isotopes. The time (in minutes) until the i-th isotope decays, is given by $X_i \sim \mathrm{Exp}(2)$, meaning that $f_X(x) = 2e^{-2x}$ for $x \geq 0$. Let Y be the number of isotopes that have decayed within the first minute (meaning $X_i \leq 1$). Give the probability mass function for Y. If it belongs to one of the families we saw in the course, it is sufficient to give that family with the correct parameter(s), e.g. $Y \sim \mathrm{Bin}(12, 0.2)$.

8. Let $X \sim \mathrm{Ber}(0.2)$ and let $Y = 1 - X^2$. Give the probability mass function for Y. If it belongs to one of the families we saw in the course, it is sufficient to give that family with the correct parameter(s), e.g. $Y \sim \mathrm{Bin}(12, 0.2)$.

9. Consider a uniform distribution on the disk with center the origin and radius 1. This means that (X,Y) can only be such that $X^2 + Y^2 \leq 1$ and the joint probability density function of X and Y is a constant on this disk. What is the value of this constant?

10. Compute $F_{X,Y}(0,1)$ for the random variables from the previous question.

Open questions

11. (a) Let X be a continuous random variable that takes values on $[0,4]$. On this interval its density function is given by $f_X(x) = \frac{3}{16}\sqrt{x}$. Now consider $Y = e^X$. Find the density function of Y. If your function description is only valid on a certain interval, then do mention this.

    (b) Consider $E[Y]$ where Y is given in part (a). It can be expressed as an integral:

$$E[Y] = \int_0^4 h(x)dx.$$

    Give $h(x)$ in this particular case. You do not need to evaluate the integral. You can solve this without having solved part (a).

    (c) $E[Y]$ is (approximately) given by one of the following values. Which one? Motivate your answer! Again, you can solve this without having solved part (a) or part (b).

    (A) $e^{1.78}$ **(B)** $e^{2.17}$ **(C)** $e^{2.33}$ **(D)** $e^{2.85}$ **(E)** $e^{4.21}$

12. I throw three fair, standard dice. Let X denote the number of fives and Y the number of sixes. As an example, if I throw (5,1,5) then $X = 2$ and $Y = 0$.

    (a) Without doing any computations, can you see whether $\text{Cov}(X,Y)$ will be negative, zero, or positive? Motivate why you think so.

    (b) Compute this covariance. If you think it is zero, the motivation from part (a) is sufficient.

    Hint: it is not necessary to construct the full joint probability mass function of X and Y.

Solutions

1. **(F)** $P(G \cap E) = P(G|E)P(E) = 0.6 \cdot 0.35 = 0.21$. Since $P(G \cup E) = P(E) + P(G) - P(G \cap E)$ we find $P(G) = 0.8 - 0.35 + 0.21 = 0.66$.

2. **(D)** Denoting by C the event that the animal has been correctly identified, and by M, F, B the events that it is a mammal, fish, bird, respectively, Bayes' rule tells us that

$$\begin{aligned} P(M|C^c) &= \frac{P(C^c|M)P(M)}{P(C^c|M)P(M) + P(C^c|F)P(F) + P(C^c|B)P(B)} \\ &= \frac{0.2 \cdot 0.6}{0.2 \cdot 0.6 + 0.5 \cdot 0.3 + 0.4 \cdot 0.1} = \frac{12}{31}. \end{aligned}$$

3. **(A)** Applying Chebyshev's inequality to the given variable X, we get

$$P\left(\left|X - \frac{1}{3}\right| \geq \frac{1}{2}\right) \leq \frac{\frac{1}{18}}{\left(\frac{1}{2}\right)^2} = \frac{4}{18}.$$

Here we chose $a = \frac{1}{2}$ since with this choice the event at the left hand side can be rewritten as $X \leq -\frac{1}{6}$ or $X \geq \frac{5}{6}$. The first option is impossible since X takes only non-negative values.

4. **(A)** By the Central Limit Theorem, the sum S approximately has a normal distribution with parameters $E[S] = 100 \cdot \frac{1}{3}$ and $\text{Var}(S) = 100 \cdot \frac{1}{18}$. Standardization then leads to

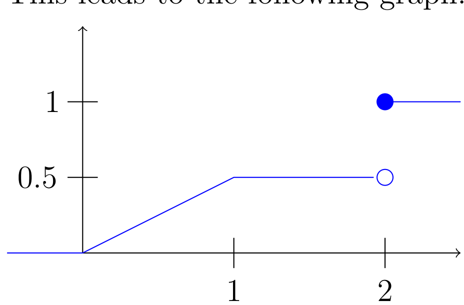
$$P(S \geq 36) = P\left(\frac{S - \frac{100}{3}}{\sqrt{\frac{100}{18}}} \geq \frac{36 - \frac{100}{3}}{\sqrt{\frac{100}{18}}}\right) \approx P(Z \geq 1.13) \approx 0.1292$$

by the Z-table.

5. **(E)**

$$E[X] = \int_1^2 \int_1^2 x \cdot \frac{3}{14}(x^2 + y^2)dxdy = \frac{3}{14} \int_1^2 \left(\frac{15}{4} + \frac{3}{2}y^2\right) dy = \frac{3}{14}\left(\frac{15}{4} + \frac{7}{2}\right) = \frac{87}{56}.$$

6. X is a mixture of a discrete and a continuous random variable, but the distribution function always has the same definition: $F_X(x) = P(X \leq x)$. Obviously, for $x < 0$ we have $F_X(x) = 0$. For $0 \leq x \leq 1$ we have $F_X(x) = P(\text{Tails}) \cdot P(U \leq x) = \frac{1}{2} \cdot x$. For $1 \leq x < 2$ the distribution function is constant, since no outcomes between 1 and 2 are possible. At $x = 2$ there is a jump, since $P(X \leq 2) = P(\text{Heads}) + P(\text{Tails}) = 1$. This leads to the following graph.



7. For each isotope, the probability that it decays in the first minute is given by $P(X_i \leq 1) = F_{X_i}(1) = 1 - e^{-2 \cdot 1} = 1 - e^{-2}$. If we call this a success, then Y counts the number of successes in 10 independent Bernoulli trials, meaning that Y has a binomial distribution, specifically $Y \sim \text{Bin}(10, 1 - e^{-2})$ or $Y \sim \text{Bin}(10, 0.86)$.

8. If $X = 0$ then $Y = 1$, if $X = 1$ then $Y = 0$. Hence Y is also a Bernoulli random variable, but success has become failure and vice versa: $Y \sim \text{Ber}(0.8)$.

9. The area of the disk D is $\pi \cdot 1^2 = \pi$. Since the joint pdf has to be a constant, and since the total volume below the graph has to be equal to 1, the joint pdf has to be given by $f_{X,Y}(x, y) = \frac{1}{\pi}$.

10. $F_{X,Y}(0, 1) = P(X \leq 0 \text{ and } Y \leq 1) = \frac{1}{2}$ since this region is exactly the left half of the disk and since the joint pdf is constant.

11. (a) Since X takes values on $[0, 4]$, $Y = e^X$ takes values on $[e^0, e^4]$. Outside this interval the density of Y is zero. On the interval we can compute it by considering the distribution function: for $0 \leq x \leq 4$,

$$F_X(x) = \int_0^x \frac{3}{16}\sqrt{t}dt = \frac{1}{8}x^{3/2}.$$

Hence, for $1 \leq y \leq e^4$,

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq \ln(y)) = F_X(\ln(y)) = \frac{1}{8}(\ln(y))^{3/2}.$$

Differentiate to find the pdf of Y:

$$f_Y(y) = \frac{3}{16}\frac{\sqrt{\ln(y)}}{y}.$$

(b)

$$E[Y] = E[e^X] = \int_0^4 e^x f_X(x)dx = \int_0^4 e^x \cdot \frac{3}{16}\sqrt{x}dx.$$

(c) In $Y = g(X) = e^X$, this g is a convex function (well-known graph of the exponential, or compute the second derivative which is the function itself, which is positive). By Jensen's inequality, we then have that

$$E[Y] = E[g(X)] \geq g(E[X]) = e^{E[X]} = e^{2.4}$$

since

$$E[X] = \int_0^4 x \cdot \frac{3}{16}\sqrt{x}dx = \frac{12}{5} = 2.4.$$

This leaves only options (D) and (E), however Y only takes values on $[1, e^4]$, so its expectation cannot be greater than e^4.

12. (a) The more fives, the fewer dice remain that can be a six. So if X is large, the probability that Y is also large decreases. This gives a negative covariance.

(b) $\text{Cov}(X,Y) = E[XY] - E[X]E[Y]$. It is straightforward to compute that $E[X] = E[Y] = \frac{1}{2}$ (they both have a Bin(3, 1/6) distribution). For $E[XY]$ we need to compute $xyP(X = x, Y = y)$ for all possible outcomes (x, y) and add these. Clearly we get zero if either x or y is zero, so we only get

- $X = Y = 1$: we need a five, a six, and one other number z (four options), which can be ordered in six different ways:
  $(5, 6, z), (6, 5, z), (5, z, 6), (6, z, 5), (z, 5, 6), (z, 6, 5)$. Since there are $6^3 = 216$ possible outcomes, this term has contribution $1 \cdot 1 \cdot 4 \cdot \frac{6}{216} = \frac{1}{9}$.
- $X = 1, Y = 2$: one five and two sixes, can be ordered in three ways: $(5, 6, 6), (6, 5, 6), (6, 6, 5)$. This term has contribution $1 \cdot 2 \cdot \frac{3}{216} = \frac{1}{36}$.
- $X = 2, Y = 1$: similar to the previous case.

This leads to $E[XY] = \frac{1}{9} + \frac{1}{36} + \frac{1}{36} = \frac{1}{6}$ and hence $\text{Cov}(X,Y) = \frac{1}{6} - \frac{1}{2} \cdot \frac{1}{2} = -\frac{1}{12}$, confirming the result from part (a).

UPDATE: a quicker way is to notice that $X + Y$ counts the number of (fives or sixes) and hence it has a Bin(3, 1/3) distribution. As mentioned before, both X and Y have a Bin(3, 1/6) distribution. The variance of a binomial distribution is given by $np(1 - p)$ so by the defining property of Covariance, we find that

$$\begin{aligned} & \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\ \text{Cov}(X,Y) \\ \Rightarrow\quad & 3 \cdot \frac{1}{3} \cdot \frac{2}{3} = 3 \cdot \frac{1}{6} \cdot \frac{5}{6} + 3 \cdot \frac{1}{6} \cdot \frac{5}{6} + 2\ \text{Cov}(X,Y) \\ \Rightarrow\quad & \text{Cov}(X,Y) = -\frac{1}{12}. \end{aligned}$$