

MATH1103 FALL 2022
PROBLEM SET 6 SOLUTIONS

This problem set is due on Wednesday, October 12 at 11:59 pm. Each problem part is worth 3 points. Collaboration is encouraged. In all cases, you must write your own solutions, and and you must cite collaborators and resources used.

Problem 1. Some exercises.

- (a) (Strang 8.4.3) Why is $p(x) = e^{-2x}$ not an acceptable probability density for $x \geq 0$? Why is $p(x) = 4e^{-2x} - e^{-x}$ (also for $x \geq 0$) not acceptable?

Solution

The function $p(x) = e^{-2x}$ is not a valid probability density function because it does not have total area 1 from 0 to ∞ . The total area is in fact

$$\int_0^{\infty} e^{-2x} dx = -\frac{1}{2}e^{-2x} \Big|_0^{\infty} = \frac{1}{2}.$$

The function $p(x) = 4e^{-2x} - e^{-x}$ is not a valid probability density function because it attains some negative values. More precisely, for all $x > \ln 4$, we have $0 < e^{-x} < \frac{1}{4}$, so that $4e^{-2x} - e^{-x} = e^{-x}(4e^{-x} - 1) < 0$, being the product of a positive number and a negative number.

- (b) (Strang 8.4.5) If $p(x) = e^{-x}$ for $x \geq 0$, find the probability that $X \geq 2$ and the approximate probability that $1 \leq X \leq 1.01$.

Solution

The probability that $X \geq 2$ is

$$\int_2^{\infty} e^{-x} dx = -e^{-x} \Big|_2^{\infty} = e^{-2}.$$

The probability that $1 \leq X \leq 1.01$ can be approximated by the area of a rectangle with width 0.01 and height $p(1) = e^{-1}$, which comes out to about 0.00368 or about 1 in 272.

We could also find the exact probability, although it is unnecessary: $\int_1^{1.01} p(x) dx$ comes out to 0.00366 or about 1 in 273. The approximation is extremely close, which suggests that the probability density function is not changing too much between 1 and 1.01!

- (c) (Strang 8.4.12) Find the mean of the distribution given by the PDF $p(x) = e^{-x}$ for $x \geq 0$. (Integrate by parts.)

Solution

The mean is given by the integral

$$\int_0^{\infty} x e^{-x} dx.$$

We integrate $x e^{-x}$ by parts using $u = x$ and $v = -e^{-x}$ to find that $\int x e^{-x} = (-x - 1)e^{-x} + C$. Therefore,

$$\int_0^{\infty} x e^{-x} dx = (-x - 1)e^{-x} \Big|_0^{\infty} = 0 - (-0 - 1) = 1.$$

- (d) (Strang 8.4.19) Supernovas are expected about every 100 years. What is the probability that you will be alive for the next one?

Hint: You can use a Poisson model with $\lambda = 0.01 \times (\text{your lifetime})$ and estimate your lifetime. You can also use an exponential distribution.

(Supernovas actually occurred in 1054 (Crab nebula), 1572, 1604, and 1987. But the future distribution doesn't depend on the date of the last one.)

Remark: Strang discusses the Poisson distribution on pages 331–332, including in the table at the end of page 332.

Solution

Let's say our lifetime is 80 years from today. If we use the Poisson model with mean expected occurrences within our lifetimes being $\lambda = 0.01 \cdot 80 = 0.8$, the model says that the probability of no occurrences is equal to $\frac{1}{0!} e^{-0.8} \lambda^0 = e^{-0.8}$. So the probability of at least one occurrence is $1 - e^{-0.8} \approx 55\%$.

If we use the exponential distribution instead, we will be modeling the expected time until the next supernova by the probability distribution $p(x) = a e^{-ax}$ with $a = 0.01$. The probability that the time until the next supernova is at most 80 years (our lifetime) is therefore

$$\int_0^{80} 0.01 e^{-0.01x} dx = -e^{-0.01x} \Big|_0^{80} = 1 - e^{-0.8} \approx 55\%,$$

same as before.

- (e) (Strang 8.4.24) What is the variance of the uniform distribution on $[0, 1]$? What is the standard deviation?

Solution

The PDF of the uniform distribution is the constant function 1 on $[0, 1]$ (and 0 elsewhere). The mean is 0.5. The variance of this distribution is therefore equal to

$$\begin{aligned}\int_0^1 \left(x - \frac{1}{2}\right)^2 dx &= \frac{1}{3} \left(x - \frac{1}{2}\right)^3 \Big|_0^1 \\ &= \frac{1}{3} \cdot \frac{1}{8} - \frac{1}{3} \cdot \left(-\frac{1}{8}\right) \\ &= \frac{1}{12}.\end{aligned}$$

The standard deviation is the square root of the variance, so it is $\frac{1}{\sqrt{12}} = \frac{1}{2\sqrt{3}}$. You could simplify this further to $\frac{\sqrt{3}}{6}$, if you like.

- (f) (Strang 8.4.33) Suppose grades have a normal distribution with mean 70 and standard deviation 10. If 300 students take the test and passing is 55, how many are expected to fail? Write your answer as a definite integral then give an approximation. What passing grade will fail 1/10 of the class?

Solution

The normal distribution with mean 70 and standard deviation 10 has probability density function

$$p(x) = \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2}.$$

The fraction of students expected to fail is the fraction of students who obtain a score less than 55, which is

$$\int_{-\infty}^{55} p(x) dx \approx 6.68\%.$$

This was numerically calculated by computer. You could also use the range 0 to 55 and get

$$\int_0^{55} p(x) dx \approx 6.68\%.$$

In fact the two numbers agree to at least 6 significant digits. In either case, the number of students expected to fail is $6.68\% \cdot 300 \approx 20$ students.

To determine the passing grade that fails 1/10 of the class, one must find the value of a for which

$$\int_{-\infty}^a p(x) dx = 0.1$$

This again has to be done by computer, either by iterative guess and check or by a z -table. Either way, one finds that a passing grade of about 57.2 will do the job.

Problem 2. You know from class that $\text{Var}(X)$ is defined as $\mathbb{E}[(X - \mu)^2]$, where $\mu = \mathbb{E}[X]$. Use calculus to prove the following famous identity for variance:

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

Hint: If you are stuck, Exercise 8.4.23 in Strang has some spoilers to get you started, but you still need to explain the step Strang asks to explain.

Solution

Let $p(x)$ be the PDF of X . We use algebra and find:

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mu)^2] \\ &= \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx \\ &= \int_{-\infty}^{\infty} (x^2 - 2x\mu + \mu^2) p(x) dx \\ &= \int_{-\infty}^{\infty} x^2 p(x) dx - 2\mu \int_{-\infty}^{\infty} x p(x) dx + \mu^2 \int_{-\infty}^{\infty} p(x) dx. \end{aligned}$$

The first term equals $\mathbb{E}[X^2]$, the second term equals $2\mu\mathbb{E}[X] = 2\mathbb{E}[X]^2$, and the third term is $\mu^2 \cdot 1 = \mathbb{E}[X]^2$ because $\int_{-\infty}^{\infty} p(x) dx = 1$. Therefore,

$$\text{Var}(X) = \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2,$$

as desired.

Problem 3. Two random numbers X and Y are drawn (independently) from a uniform distribution on $[0, 1]$. Find $\mathbb{E}[\max(X, Y)]$.

Hint: This is a pretty fun problem. Start by finding the CDF of $\max(X, Y)$. Use the fact that $\max(X, Y) \leq z$ if and only if both $X \leq z$ and $Y \leq z$.

Optional challenge: Generalize to the $\max(X_1, \dots, X_n)$ where each X_i is drawn from a uniform distribution on $[0, 1]$.

Solution

Here is the intended solution. At the end I will show a very sleek solution that will amaze you. Let z be an arbitrary real number. Since $\max(X, Y) \leq z$ if and only if

$X \leq z$ and $Y \leq z$, we see that

$$\Pr[\max(X, Y) \leq z] = \Pr[X \leq z] \cdot \Pr[Y \leq z] = z \cdot z = z^2.$$

Let Z denote the random variable $\max(X, Y)$. Since $\text{CDF}_Z(z)$ is the same thing as $\Pr[\max(X, Y) \leq z]$, we find that $\text{CDF}_Z(z) = z^2$. Therefore, $\text{PDF}_Z(z) = \frac{d}{dz} z^2 = 2z$. Hence

$$\mathbb{E}[\max(X, Y)] = \mathbb{E}[Z] = \int_0^1 z \text{PDF}_Z(z) dz = \int_0^1 2z^2 dz = \frac{2}{3}.$$

The same method works for the optional challenge, just now $\text{CDF}_Z(z)$ is the probability that n independent random numbers are all less than z , which is z^n .

Solution

Here is the sleek solution. We pick 3 random numbers in $[0, 1]$, calling the first two X and Y respectively and calling the last number Z . For any values of X and Y , the probability that Z is less than at least one of X or Y is the same as the value of $\max(X, Y)$, because Z is random in $[0, 1]$. Therefore, $\mathbb{E}[\max(X, Y)]$ is the same as $\Pr[Z < X \text{ or } Z < Y]$. This probability is just the probability that Z is not the largest of the 3 numbers X, Y, Z . The probability that Z is the largest of the 3 numbers X, Y, Z is $1/3$, by symmetry. Therefore, the probability that Z is not the largest is $2/3$.

To generalize to the optional challenge, the answer now is just the probability that Z is not the largest among $n + 1$ numbers. The probability of this is $n/(n + 1)$ by the same reasoning. Very cool!

Problem 4 (Optional). [This video](#) talks about the surprisingly non-intuitive problem of generating a random point inside a circle. After watching the video, think about the following.

- What do you think is the fairest way to pick a random point on a circle?
- The video mentioned something called inverse transform sampling. To test if you understood it, what is the probability density function of X^2 if X is drawn from a uniform distribution on $[0, 1]$? (Recall we saw in class a while ago that the average value of X^2 was $1/3$.)

Hint: As usual, start by finding the CDF.

Spoiler: the PDF of X^2 is $\frac{1}{2\sqrt{x}}$ for $0 \leq x \leq 1$.

Problem 5 (Optional).

- (a) Let X and Y be two independent random variables with PDFs $p(x)$ and $q(x)$ respectively. Prove that the PDF of $X + Y$ is given by

$$\text{PDF}_{X+Y}(x) = \int_{-\infty}^{\infty} p(t)q(x-t) dt.$$

Hint: Find the CDF of $X + Y$, then differentiate under the integral sign.

- (b) We write $X \sim N(\mu, \sigma)$ to mean that X follows a normal distribution with mean μ and standard deviation σ . Using part (a), prove that if $X \sim N(0, 1)$ and $Y \sim N(0, 1)$, then $X + Y \sim N(0, \sqrt{2})$.

Solution

Let me know if you want to see solutions to these optional problems!