

STAT 310: Homework #3

Due on October 6, 2017 at 5:00 PM

Guerra

Joel Abraham

Problem 1

RT 2.5.5. The cumulative distribution function $F(x)$ of a random variable X is given by

$$F(x) = \begin{cases} 0, & -\infty < x < -1 \\ 0.2, & -1 \leq x < 3 \\ 0.8, & 3 \leq x < 9 \\ 1, & x \geq 9 \end{cases}$$

Write down the values of the random variable X and the corresponding probabilities $p(x)$. Also find $\mathbb{E}[X]$ and $\text{SD}(X)$.

Solution

$$p(x) = \begin{cases} 0.2, & x = -1 \\ 0.6, & x = 3 \\ 0.2, & x = 9 \\ 0, & \text{otherwise} \end{cases}$$

. Then, $\mathbb{E}[X] = (-1)(0.2) + (3)(0.6) + (9)(0.2) = -0.2 + 1.8 + 1.8 = 3.4$. $\text{Var}(X) = \sigma^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (1)(0.2) + (9)(0.6) + (81)(0.2) + 3.4^2 = 0.2 + 5.4 + 16.2 - 11.56 = 10.24 \implies \text{SD}(X) = \sqrt{10.24} = 3.2$.

Problem 2

Rose is planning a dinner party for seven guests. All her invited friends will be assistant professors, all of whom are working 16 hours a day to get tenure. Being so busy with their careers, each professor will accept the invitation according to the flip of a fair coin: heads accept, tails don't accept the invitation. Rose also knows that each professor that accepts the invitation will actually attend with chance 80%. Rose decides to invite 12 professors. The professors have almost no social life so they will make decisions independently of each other. What is the probability that Rose will end up with more professors than she can accommodate?

Solution

The probability of a professor attending is $P(A) = (0.5)(0.8) = 0.4$, where A is the event that a professor attends. Thus, we want to find the probability of 8 or more professors attending. Since this follows a binomial distribution, we have (where M is the event that Rose ends up with more professors than she can accommodate)

$$\begin{aligned} P(M) &= \sum_{k=8}^{12} \binom{12}{k} (0.4)^k (0.6)^{12-k} \implies \\ P(M) &= \binom{12}{8} (0.4)^8 (0.6)^4 + \binom{12}{9} (0.4)^9 (0.6)^3 + \binom{12}{10} (0.4)^{10} (0.6)^2 + \binom{12}{11} (0.4)^{11} (0.6)^1 + \binom{12}{12} (0.4)^{12} \implies \\ P(M) &= 0.0573 \end{aligned}$$

Problem 3

ACT and SAT college entrance exams are taken by thousands of students each year. The mathematics scores on each exam are normally distributed. In recent years, the SAT math scores have had a mean of 480 and

SD of 100, while the ACT mean and SD are 18 and 6, respectively. For purposes of this exercise, assume SAT scores range from 400 to 800 and ACT scores range from 1 to 36.

- A certain engineering program has a minimum SAT math score of 600 for admission. What percentage of students are eligible for the program?
- What is a comparable ACT standard for admission to the engineering program?
- The engineering program accepts 30 students per year. How many of these 30 students are expected to have an SAT math score of 750 or higher?

Solution

- The probability density function for the normally distributed SAT math scores is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

. Thus, for $\mu = 480$ and $\sigma = 100$, we have

$$f(x) = \frac{1}{100\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-480}{100}\right)^2}$$

. We now consider

$$F(x) = \int_{-\infty}^x f(x)dx = \int_{-\infty}^x \frac{1}{100\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-480}{100}\right)^2} dx$$

$$P(X \geq 600) = 1 - P(X < 600) = 1 - F(600) = \int_{600}^{\infty} \frac{1}{100\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-480}{100}\right)^2} dx = 0.115.$$

We compute this using the R command: `pnorm(600, mean=480, sd=100, lower.tail = FALSE) = 0.115`.

- A comparable ACT score is given by: `qnorm(0.1150697, mean=18, sd=6, lower.tail = FALSE) = 25.2 \approx 25`
- Let A represent the event that a student has an SAT math score greater than or equal to 750 and let B represent the event that a student has an SAT math score greater than or equal to 600 (is accepted into the engineering program). Now, $P(A) = \text{pnorm}(750, \text{mean}=480, \text{sd}=100, \text{lower.tail} = \text{FALSE}) = 0.00347$ and $P(B) = 0.115$. Then,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.00347}{0.115} = 0.0301.$$

Thus, we would expect $0.0301 * 30 = 0.904 \approx 1$ student in the engineering program to score 750 or higher.

Problem 4

A jury of 12 persons is to be selected from a pool of 25 persons called to jury duty. The pool of 25 persons includes 12 retired persons, 6 employed persons, 2 unemployed and 5 students. Assume that each person is equally likely to be selected.

- What is the probability that two students will be selected?
- What is the probability that at least half the jury will be comprised of retired persons?
- What is the probability that the jury will contain exactly twice as many retired persons as employed persons?

Solution

- (a) Let T represent the event that exactly two students are selected. Then, $P(T)$ is the number of ways we can choose a jury with two students divided by the number of total juries. This is given by

$$P(T) = \frac{\binom{5}{2}\binom{20}{10}}{\binom{25}{12}} = 0.355$$

- (b) The probability that at least half the jury is composed of retired persons (which we label event R) is the probability that the jury has 6 - 12 retired persons. Then, we have

$$P(R) = \sum_{k=6}^{12} \frac{\binom{12}{k}\binom{13}{12-k}}{\binom{25}{12}} = \frac{\binom{12}{6}\binom{13}{6}}{\binom{25}{12}} + \frac{\binom{12}{7}\binom{13}{5}}{\binom{25}{12}} + \frac{\binom{12}{8}\binom{13}{4}}{\binom{25}{12}} + \dots + \frac{1}{\binom{25}{12}} = 0.582$$

- (c) Let E be the event that the jury contains exactly twice as many retired people as employed people. Then, the jury can have 2, 3, or 4 employed people, and 2, 4, 6, or 8 retired people, respectively. Thus,

$$P(E) = \sum_{k=2}^4 \frac{\binom{6}{k}\binom{12}{2k}\binom{7}{12-3k}}{\binom{25}{12}} = \frac{\binom{6}{2}\binom{12}{4}\binom{7}{6}}{\binom{25}{12}} + \frac{\binom{6}{3}\binom{12}{6}\binom{7}{3}}{\binom{25}{12}} + \frac{\binom{6}{4}\binom{12}{8}}{\binom{25}{12}} = 0.136$$

Problem 5

Let $Y \sim \text{Exp}(\beta)$. Show that Y has a memoryless property. Thus, if $s \geq 0$ and $t \geq 0$, then

$$P(Y > s + t | Y > s) = P(Y > t)$$

Solution

Since $Y \sim \text{Exp}(\beta)$, we have $f(x) = \frac{1}{\beta}e^{-\frac{x}{\beta}}$. And, $P(Y > s + t | Y > s) = \frac{P(Y > s+t)}{P(Y > s)}$. First, note that:

$$P(Y > a) = \int_a^{\infty} f(x)dx = \int_a^{\infty} \frac{1}{\beta}e^{-\frac{x}{\beta}}dx = \left[-e^{-\frac{x}{\beta}}\right]_a^{\infty} = -e^{-\frac{\infty}{\beta}} + e^{-\frac{a}{\beta}} = e^{-\frac{a}{\beta}}.$$

Thus, we have

$$\frac{P(Y > s + t)}{P(Y > s)} = \frac{e^{-\frac{s+t}{\beta}}}{e^{-\frac{s}{\beta}}} = e^{-\frac{s+t}{\beta} + \frac{s}{\beta}} = e^{-\frac{s+t-s}{\beta}} = e^{-\frac{t}{\beta}} = P(Y > t).$$

Problem 6

The number of arrivals N at a supermarket checkout counter in the time interval from 0 to t follows a Poisson distribution with mean λt . Let T denote the length of time until the first arrival. Find the probability density function of T . Hint: Consider $P(T > t_0)$.

Solution

Since we know that arrivals follow a Poisson distribution, we have

$$P(N = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}.$$

Now, we seek to find

$$F(t) = P(T < t) = 1 - P(T > t).$$

Since T is the length of time until the first arrival, $P(T > t)$ is the probability that the first arrival occurs after some time t , which is equivalent to the probability that no arrivals occur between time 0 and t . Thus,

$$P(T > t) = P(N = 0) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t} \implies F(t) = P(T < t) = 1 - e^{-\lambda t}.$$

Since the probability density function is the derivative of the cumulative distribution function, we have

$$f(t) = \frac{d}{dt}F(t) = \lambda e^{-\lambda t}.$$

Problem 7

Let Y be a continuous random variable with pdf, $f(y)$, and cdf, $F(y)$. We are often interested in conditional probabilities of the form $P(Y \leq y | Y \geq c)$. For example, the probability that cell phone battery will last less than 3 years, given that it is now 1 year old.

- (a) Show that for $y \geq c$,

$$F_c(y) = P(Y \leq y | Y \geq c) = \frac{F(y) - F(c)}{1 - F(c)}$$

- (b) Show that the above function satisfies all the properties of a distribution function. (a) $F_c(y)$ approaches 0 as $y \implies -\infty$. (b) $F_c(y)$ approaches 1 as $y \implies \infty$. (c) $F_c(y)$ is non-decreasing; for $y_1 < y_2$, then $F_c(y_1) \leq F_c(y_2)$.

- (c) Since $F_c(y)$ is a distribution function of a continuous r.v., then it has a pdf. Show that the pdf is

$$f_c(y) = \frac{f(y)}{1 - F(c)}, y \geq c.$$

- (d) Suppose you are waiting for a call back from a job interview. Assume the waiting time for the call back, Y , follows an exponential distribution with mean 24 hours.

- (i) What is chance that you will get the call within 24 hours?
- (ii) What is the expected value of Y ?
- (iii) Given that 8 hours have already passed, what is the chance that you will get the call within 24 hours?
- (iv) Given that 8 hours have already passed, what is the expected value of Y ?

Solution

- (a) For $y \geq c$,

$$P(Y \leq y | Y \geq c) = \frac{P(c \leq Y \leq y)}{P(Y \geq c)}.$$

$$P(Y \geq c) = 1 - P(Y < c) = 1 - F(c), \text{ and}$$

$$P(Y \leq y | Y \geq c) = \int_c^y f(x)dx = \int_c^{-\infty} f(x)dx + \int_{-\infty}^y f(x)dx = F(y) - F(c).$$

Thus, we have

$$F_c(y) = P(Y \leq y | Y \geq c) = \frac{F(y) - F(c)}{1 - F(c)}$$

- (b) Since F is a distribution function, we know $\lim_{x \rightarrow \infty} F(x) = 1$ and $\lim_{x \rightarrow -\infty} F(x) = 0$, thus

$$\lim_{y \rightarrow \infty} F_c(y) = \lim_{y \rightarrow \infty} \frac{F(y) - F(c)}{1 - F(c)} = \frac{\lim_{y \rightarrow \infty} F(y) - F(c)}{1 - F(c)} = \frac{1 - F(c)}{1 - F(c)} = 1.$$

For the case when y approaches $-\infty$, we have $y < c$, thus $F_c(y) = 0$ since $P(Y \leq y | Y \geq c) = 0$ because when $y < c$, we have $P(Y \leq y \cap Y \geq c) = P(Y < c \cap Y \geq c) = P(\emptyset) = 0$. Now, we show that $F_c(y)$ is nondecreasing since $F(x)$ is nondecreasing. For $y_1 < y_2$, we have

$$F(y_1) \leq F(y_2) \implies F(y_1) - F(c) \leq F(y_2) - F(c) \implies \frac{F(y_1) - F(c)}{1 - F(c)} \leq \frac{F(y_2) - F(c)}{1 - F(c)} \implies F_c(y_1) \leq F_c(y_2).$$

- (c) The probability distribution function is the derivative of the cumulative distribution, so we first observe that

$$F_c(y) = \frac{F(y)}{1 - F(c)} - \frac{F(c)}{1 - F(c)},$$

thus since $(1 - F(c))^{-1}$ is a constant,

$$f_c(y) = \frac{d}{dy} F_c(y) = \frac{d}{dy} \left(\frac{F(y)}{1 - F(c)} \right) = \frac{f(y)}{1 - F(c)}, y \geq c$$

- (d) $Y \sim \text{Exp}(\beta)$, where $\beta = 24$

- (i) The chance of a callback within 24 hours is equal to

$$F(24) = \int_0^{24} \frac{1}{24} e^{-\frac{x}{24}} dx = \left[1 - e^{-\frac{x}{24}} \right]_0^{24} = 1 - \frac{1}{e}$$

- (ii) The expected value is equivalent to the mean, thus $\mathbb{E}[Y] = \beta = 24$

- (iii) We use the definition of $F_8(y)$ as follows. We must first calculate $F(8) = 1 - e^{-\frac{1}{3}}$, thus

$$F_8(24) = \frac{F(24) - F(8)}{1 - F(8)} = \frac{1 - e^{-1} - (1 - e^{-\frac{1}{3}})}{1 - (1 - e^{-\frac{1}{3}})} = \frac{e^{-\frac{1}{3}} - e^{-1}}{e^{-\frac{1}{3}}} = 1 - e^{(-1 + \frac{1}{3})} = 1 - e^{-\frac{2}{3}}$$

- (iv) Given that 8 hours have passed, we define

$$F_8(y) = 1 - e^{(\frac{8-y}{24})}.$$

Thus, by linearity of expectation and part (b), we have

$$\mathbb{E}[Y] = \int_0^\infty y f_8(y) dy = \int_0^\infty \frac{y f(y)}{1 - F(8)} dy = \frac{1}{1 - F(8)} \int_0^\infty y f(y) dy = e^{-\frac{1}{3}} \int_0^\infty y f(y) dy = 24 e^{-\frac{1}{3}} = 17.197$$

Problem 8

Problem 3.4. Severe snow storm is defined as a storm whose snowfall exceeds 10 inches. Let X be the amount of snowfall in a severe snow storm. The cumulative distribution function (CDF) of X in a given town is:

$$F_x(x) = \begin{cases} 1 - \left(\frac{10}{x}\right)^4 & x \geq 10 \\ 0 & x < 10 \end{cases}$$

- (a) Determine the median of X .

- (b) What is the expected amount of snowfall in the town in a severe snow storm?
- (c) Suppose a disastrous snow storm is defined as a storm with over 15 inches of snowfall. What percentage of the severe snow storms are disastrous?
- (d) Suppose the probability that the town will experience 0, 1, and 2 severe snow storms in a year is 0.5, 0.4, and 0.1, respectively. Determine the probability that the town will not experience a disastrous snow storm in a given year. Assume the amounts of snowfall between storms are statistically independent.

Solution

- (a) The median is the point at which $F_x(x) = 0.5$, thus

$$1 - \left(\frac{10}{x}\right)^4 = 0.5 \implies \left(\frac{10}{x}\right)^4 = 0.5 \implies \frac{10}{x} = (0.5)^{0.25} \implies x = \frac{10}{(0.5)^{0.25}} \implies x = 11.892.$$

- (b) The density function is given by

$$f(x) = \frac{d}{dx}(F)_x(x) = \frac{d}{dx}(1) - \frac{10^4}{x^4} = 4\left(\frac{10^4}{x^5}\right).$$

Thus, we calculate expected value via:

$$\mathbb{E}[X] = \int_{10}^{\infty} xf(x)dx = 4 * 10^4 \int_{10}^{\infty} \frac{1}{x^4}dx = -\frac{4}{3} * 10^4 \left[\frac{1}{x^3} \right]_{10}^{\infty} = \frac{4}{3} * 10^4 \frac{1}{10^3} = 13.333$$

- (c) We find the probability that a severe snow storm is disastrous by

$$P(X > 15) = 1 - P(X \leq 15) = 1 - F_x(15) = 1 - \left(1 - \left(\frac{2}{3}\right)^4\right) = \left(\frac{2}{3}\right)^4 = 0.1975$$

- (d) The probability that a severe snow storm is not disastrous is $P(X \leq 15) = 1 - 0.1975 = 0.802$. Thus, the probability of no disastrous storms is given by $P(D = 0) = 0.5 * (1) + 0.4 * (0.802) + 0.1 * (0.802)^2 = 0.885$.

Problem 9

Problem 3.12. The total load, X (in tons), on the roof of a building has the following PDF:

$$f_x(x) = \begin{cases} \frac{24}{x^3} & 3 \leq x \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Determine and plot the CDF of X , i.e., $F_x(x)$.
- (b) What is the expected total load?
- (c) Determine the coefficient of variation of the total load.
- (d) Suppose the roof can carry only 5.5 tons before collapse. What is the probability that the roof will collapse?

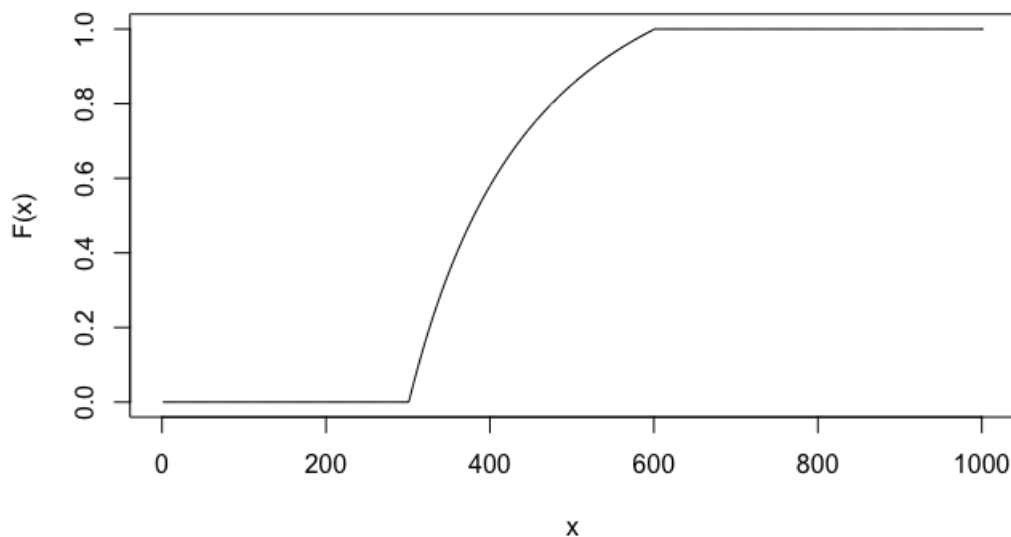
Solution

(a)

$$F_x(x) = \int_{-\infty}^x f_x(x)dx = \int_3^x f_x(x)dx = 24 \int_3^x \frac{1}{x^3}dx = -12 \left[\frac{1}{x^2} \right]_3^x = -\frac{12}{x^2} + \frac{4}{3}.$$

Thus,

$$F_x(x) = \begin{cases} 0 & x < 3 \\ -\frac{12}{x^2} + \frac{4}{3} & 3 \leq x \leq 6 \\ 1 & x > 6 \end{cases}$$

Cumulative Distribution Function

(b) The expected value of the total load is given by

$$\mathbb{E}[X] = \int_3^6 x f_x(x)dx = 24 \int_3^6 \frac{1}{x^2}dx = 24 \left[-\frac{1}{x} \right]_3^6 = 24 \left[-\frac{1}{6} + \frac{1}{3} \right] = 4$$

(c) The variance is given by

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = -\mathbb{E}[X]^2 + \int_3^6 x^2 f_x(x)dx = 24 \int_3^6 \frac{1}{x}dx - 16 = 24 [\ln x]_3^6 - 16 = 24 \ln 2 - 16$$

Thus since $\mathbb{E}[X] = 4$, we can calculate coefficient of variation as

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}[X]}}{\mathbb{E}[X]} = \frac{\sqrt{24 \ln 2 - 16}}{4} = 0.199$$

(d) The probability that the roof will collapse is the probability that the load is greater than 5.5 tons, thus

$$P(X > 5.5) = 1 - P(X \leq 5.5) = 1 - F_x(5.5) = 1 + \frac{12}{5.5^2} - \frac{4}{3} = 0.0633$$