Problem Set 3 - Discrete and Continuous Random Variables

ECO 104 (Section 7 and 9)
Instructor: Shaikh Tanvir Hossain

Due Date: No Need to Submit

§

Please Note: No need to submit, but please solve before Mid - 2 Exam

Discrete Random Variables

1. From Anderson et al. (2020) Chapter 5, #1

   Consider the experiment of tossing a coin twice.

   (a) List the experimental outcomes.

   (b) Define a random variable that represents the number of heads occurring on the two tosses.

   (c) Show what value the random variable would assume for each of the experimental outcomes.

   (d) Is this random variable discrete or continuous?

   Solution: We solved a problem with three coin tosses, this one is easier, so try this!

2. From Anderson et al. (2020) Chapter 5, #2

   Consider the experiment of a worker assembling a product.

   (a) Define a random variable X that represents the time in minutes required to assemble the product.

   (b) What values may X take?

   (c) Is X a discrete or continuous random variable?

   Solution: This is an open question, and answer can vary depending upon your thoughts about how much time does he/she need.

3. From Anderson et al. (2020) Chapter 5, #7.

   The probability mass function or PMF for the random variable X follows.

x	$f(x)$
20	0.20
25	0.15
30	0.25
35	0.40

   (a) Is this a valid PMF? Explain.

(b) What is $\mathbb{P}(X = 30)$?

(c) What is $\mathbb{P}(X \leq 25)$?

(d) What is $\mathbb{P}(X \geq 30)$?

(e) What is $\mathbb{E}(X)$ and (X)?

(f) What is $\mathbb{E}(X^2)$ and (X^3)?

(g) What is $\mathbb{E}(2X + 3)$ and $(2X + 3)$?

(h) What is $\mathbb{E}(2X^2 + 3)$ and $(2X^2 + 3)$?

Solution:

a. Yes, it is a valid PMF, becuse $f(20) + f(25) + f(30) + f(35) = 1$. This is the property written in a Theorem in the slides.

b. We can calculate this straight from PMF, $\mathbb{P}(X = 30) = f(30) = 0.25$

c. This is a cumulative probability upto 25, $\mathbb{P}(X \leq 25) = f(20)+f(25) = 0.20+0.15 = 0.35$

d. Again another cumulative probability, please do it on your own.

4. From Anderson et al. (2020) Chapter 5, #15.

The following table provides a PMF for the random variable X.

x	$f(x)$
3	0.25
6	0.50
9	0.25

(a) Compute $\mathbb{E}(X)$ and (X).

(b) Compute the standard deviation of X.

(c) What will happen to variance if we change the distribution as follows,

x	$f(x)$
3	0.3333
6	0.3333
9	0.3333

Solution:

a. We discussed the idea of Expected Value, this is explained in the slides. We can simply apply the formula

$$\begin{aligned}\mathbb{E}(X) &= (3 \times f(3)) + (6 \times f(6)) + (9 \times f(9)) \\ &= (3 \times 0.25) + (6 \times 0.50) + (9 \times 0.25) = 6\end{aligned}$$

Recall the idea of the Expected value - it is similar to the concept average. This is often called a measure for central tendency. In this case 6 is a value that X takes with maximum probability. Not always $\mathbb{E}(X)$ will be a value of the random variable X. So the interpretation of the Expectation is, this is a number around which the random variable takes its values with higher probability.

Let's calculate the variance of X now, recall variance is measure for the dispersion.

$$\begin{aligned}\mathbb{V}(X) = \mathbb{E}\left((X-6)^2\right) &= ((3-6)^2 \times f(3)) + ((6-6)^2 \times f(6)) + ((9-6)^2 \times f(9)) \\ &= 4.5\end{aligned}$$

b. The standard deviation is simply $\sqrt{4.5} = 2.12$. The idea of the standard deviation is it simply gives us a measure of dispersion in the same unit of the random variable.

c. Now the distribution has changed and all values have equal probability (note this is discrete uniform), without doing any calculation you can tell that the variance will probably increase, this is because before the the probability was maximum at 6, so we had more weights on the center, but now all values have same probability, so the values are more dispersed. If you calculate the mean, now you will see $\mathbb{E}(X) = 6$, but the variance is now 5.9994.

5. From Anderson et al. (2020) Chapter 5, #16 (slightly modified)

   The following table provides a probability distribution for the random variable Y,

y	$f(y)$
2	0.20
4	0.30
7	0.40
8	0.10

   (a) Compute $\mathbb{E}(Y)$ and (Y).
   (b) Compute $\mathbb{E}(Y^2)$.
   (c) Calculate $\mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2$. Check whether this is equal to (Y).

Solution: a. Calculating Expected value and Variance is same as the last problem, so I am skipping this here, but if you do this then you should get $\mathbb{E}(Y) = 5.2$, and $(Y) = 4.56$

b. This one is asking for calculating $\mathbb{E}(g(Y))$, when the function $g(Y) = Y^2$. In general for any function when we take expectation, the principle is same. So in this case,

$$\mathbb{E}(Y^2) = \left(2^2 \times f(2)\right) + \left(4^2 \times f(4)\right) + \left(7^2 \times f(7)\right) + \left(8^2 \times f(8)\right) = 31.6$$

In general you can do this for any $g()$, this is sometimes called LOTUS (Law of the the Unconscious Statistician), this name because technically speaking we are not using the distribution of $g(Y)$, we are using the distribution of Y, and it turns out we are fine! So we are doing something unconsciously, but there is no problem with that :)

Try to find $\mathbb{E}(3Y + 10)$, applying LOTUS. Here is another task for you, do this applying LOTUS, and you will see that $\mathbb{E}(3Y + 10) = 3\mathbb{E}(Y) + 10$. This gives us a very interesting relationship that is Expectation goes through a constant, and Expectation of a constant is the constant itself.

Now you can verify that $\mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = 31.6 - (5.2)^2 = 31.6 - 27.04 = 4.56$, Ah ha! we got the variance back, this relation will always hold! So this means we will always have

$$\mathbb{V}(Y) = \mathbb{E}\left[(Y - \mathbb{E}(Y))^2\right] = \mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2$$

Where the first equality is the definition of Variance and the second equality is another relation we just found to be true!

6. From Anderson et al. (2020) Chapter 5, #31.

   Consider a Binomial experiment with 2 trials and $p = .4$. This means we have a random variable X such that $X \sim \text{Bin}(2, 0.4)$

   (a) Compute $f(0), f(1), f(2)$ and interpret the value.

   (b) Compute the probability of at least one success, this means $\mathbb{P}(X \geq 1)$

   (c) Compute $\mathbb{E}(X)$ and (X).

Solution: In this case we need to use the mass function (or PMF) of Binomial distribution. The PMF is the following function (given in Definition 3.9, Chapter 3, page 40)

$$\begin{aligned} f(x; n, p) &= \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n \\ &= 0, \text{ otherwise} \end{aligned}$$

So let's start the calculation (finish the calculation using your calculator!)

$$f(0) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} (0.4)^0(1 - 0.4)^{2-0} = [2]0 \times 1 \times (0.6)^2 = .36$$

What is the interpretation of $f(0)$? Recall X is the random variable which represents the total number of success if we toss a coin twice. Here $f(0) = P(X = 0)$. So in this case $f(0)$ is the probability of 0 success if we toss a coin twice.

We can also give a frequency interpretation, that is if we toss a coin twice, then if we repeat this experiment 100 times, then 36 times we will get 0

Now we need $f(1)$

$$f(1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} (0.4)^1(1 - 0.4)^{2-1} = [2]1 \times (0.4) \times (0.6)^{2-1} = ...$$

Similarly interpret $f(1)$. Finally you need to calculate and interpret $f(2)$.

b) Here we need to calculate

$$\mathbb{P}(X \geq 1) = f(1) + f(2) + 0 = f(1) + f(2)$$

I gave 0 just to remind you that, for all the other points greater than 2, the function will take value 0.

c) We can directly apply the Expectation and Variance formula for the Binomial Random variable, this is given in slide 41, Chapter. In general if $X \sim \text{Bin}(n, p)$, then

$$\mathbb{E}(X) = np \quad \text{and} \quad \mathbb{V}(X) = n\text{ı}(1 - p)$$

So $\mathbb{E}(X) = np = 2 \times 0.4$ and $(X) = n\text{ı}(1 - p) = 2 \times 0.4 \times 0.6$ ... finish the calculation.

7. From Anderson et al. (2020) Chapter 5, #32.

   Consider a Binomial experiment with $n = 10$ and $p = .10$. This means we have a random variable $X \sim \text{Bin}(10, 0.10)$

   (a) Compute $f(0)$.

   (b) Compute $f(2)$.

   (c) Compute $\mathbb{P}(X \leq 2)$.

   (d) Compute $\mathbb{P}(X \geq 1)$.

   (e) Compute $\mathbb{E}(X)$ and (X).

8. From Anderson et al. (2020) Chapter 5, #33.

Consider a binomial experiment with $n = 20$ and $p = .70$. This means we have a random variable $X \sim \text{Bin}(20, 0.70)$

(a) Compute $f(12)$.

(b) Compute $f(16)$.

(c) Compute $\mathbb{P}(X \geq 16)$.

(d) Compute $\mathbb{P}(X \leq 15)$.

(e) Compute $\mathbb{E}(X)$ and (X).

9. From Anderson et al. (2020) Chapter 5, #36.

Number of Defective Parts. When a new machine is functioning properly, only 3% of the items produced are defective. Assume that we will randomly select two parts produced on the machine and that we are interested in the number of defective parts found.

(a) How can you think about a Bernoulli random variable here?

(b) How can you think about a Binomial random variable here? And what is the key condition under which we can think about a Binomial random variable?

(c) What are the possible values of the Binomial random variable?

(d) What is the mean and variance of the Binomial random variable?

Solution:

- a. You should always think about the Bernoulli random variable here first. In this case the Bernoulli random variable is *whether a part is defective or not*. So we can think about a Bernoulli random variable $X \sim \text{Bernoulli}(0.03)$. The values of the Bernoulli random variable X are 0 and 1, in this case the mean of the Bernoulli random variable is $\mathbb{E}(X) = 0.03$ and variance is $(X) = 0.03 \times (1 - 0.03)$.

- b. Now whenever we have more than one trial, and we are thinking about the total number of success, in the same experiment we can think Binomial random variable. In this case, the trial is *to select a part*, now if we select two parts, then the number of defective parts will be a Binomial random variable. So we can think about another random variable Y which follows Binomial distribution, we write $Y \sim \text{Bin}(2, 0.03)$. Here Y represents *the number of defective parts in* 2 *trials*. So Y can take values $0, 1, 2$. (We used Y to separate the Bernoulli and Binomial random variables, important these are two different random variables)

- c. The possible values of the Binomial random variable are $0, 1, 2$.

- d. The mean and variance of the Binomial random variable are $\mathbb{E}(Y) = 2 \times 0.03$ and $(Y) = 2 \times 0.03 \times (1 - 0.03)$.

10. From Anderson et al. (2020) Chapter 5, #41.

Introductory Statistics Course Withdrawals. A university found that 20% of its students withdraw without completing the introductory statistics course. Assume that 20 students registered for the course.

(a) Compute the probability that 2 or fewer will withdraw.

(b) Compute the probability that exactly 4 will withdraw.

(c) Compute the probability that more than 3 will withdraw.

(d) Compute the expected number of withdrawals.

Solution: Again we can a Bernoulli random variable whether a single student will withdraw or not, and we will repeat this experiment 20 times. then we can think about a Binomial random variable, $X \sim \text{Bin}(20, .20)$. So X represents the number of withdrawals in 20 trials (here success means withdrawal), it's clear that the value of X can be $0, 1, 2, \ldots, 20$. Now you should be able to solve the rest.

- a. $\mathbb{P}(X \leq 2) = f(0) + f(1) + f(2)$
- b. $\mathbb{P}(X = 4) = f(4)$
- c. $\mathbb{P}(X > 3) = 1 - \mathbb{P}(X \leq 3)$
- d. $\mathbb{E}(X) = 20 \times 0.20$

Use the Binomial PMF to solve this problem.

11. From Anderson et al. (2020) Chapter 5, #42.

State of the Nation Survey. Suppose a sample of 20 Americans is selected as part of a study of the state of the nation. The Americans in the sample are asked whether or not they are satisfied with the way things are going in the United States.

(a) Compute the probability that exactly 4 of the 20 Americans surveyed are satisfied with the way things are going in the United States.

(b) Compute the probability that at least 2 of the Americans surveyed are satisfied with the way things are going in the United States.

(c) For the sample of 20 Americans, compute the expected number of Americans who are satisfied with the way things are going in the United States.

(d) For the sample of 20 Americans, compute the variance and standard deviation of the number of Americans who are satisfied with the way things are going in the United States.

Solution: Here we can think about a Binomial random variable, $X \sim \text{Bin}(20, .5)$. So X represents the number of Americans who are satisfied with the way things are going in the United States if we ask 20 citizens. It's clear that the value of X can be $0, 1, 2, \ldots, 20$. Here we didn't have any p for Binomial, we are assuming $p = 0.5$, this is because we don't know the actual p, but we are assuming that the probability of being satisfied or not is same for all citizens.

Continuous Random Variables

12. Suppose following is a density function for a random variable X that takes any value between 0 and 1

$$f(x) = 3x^2, \text{ where } 0 \leq x \leq 1$$

(a) Is this is a valid density function?

(b) Calculate $\mathbb{P}(X < 0.5)$

(c) Calculate $\mathbb{P}(X > 0.6)$

(d) Calculate $\mathbb{E}(X)$ and $\mathbb{V}(X)$

Solution:

(a) To check if this is a valid density function, we need to check two things,

1. $f(x) \geq 0$ for all x in the support of X, which is $[0, 1]$.
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

For the first condition, we can see that $f(x) = 3x^2 \geq 0$ for all $x \in [0, 1]$. So this condition is satisfied.

For the second condition, we can calculate the integral,

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_0^1 3x^2 dx \\ &= 3\left[\frac{x^3}{3}\right]_0^1 \\ &= [x^3]_0^1 = 1^3 - 0^3 = 1\end{aligned}$$

13. From (Anderson et al. 2020) Chapter 6, #1

The random variable X is known to be uniformly distributed between 1.0 and 1.5 (**in notation we write** $X \sim$ **Unif**$(1, 1.5)$, they both mean same thing!)

(a) Show the graph of the probability density function.

(b) Compute $\mathbb{P}(X = 1.25)$.

(c) Compute $\mathbb{P}(1.0 \leq X \leq 1.25)$.

(d) Compute $\mathbb{P}(1.20 < X < 1.5)$.

Solution:

(a) The density function of X is

$$f(x) = \frac{1}{b-a} = \frac{1}{1.5-1.0} = \frac{1}{0.5} = 2 \quad \text{for } 1.0 \leq x \leq 1.5$$

This is a simple plot, we did the plot in the class

(b) X is a continuous random variable, so this means

$$\mathbb{P}(X = 1.25) = 0$$

this holds for any value, not just 1.25

(c) There are two ways to do this, the hard way is,

$$\begin{aligned}\mathbb{P}(1.0 \leq X \leq 1.25) &= \int_{1.0}^{1.25} f(x)dx \\ &= \int_{1.0}^{1.25} 2dx \\ &= 2[x]_{1.0}^{1.25} = 2(1.25-1.0) = 2(0.25) = 0.5\end{aligned}$$

the easy way is drawing the picture of the density function, and then applying the formula for the area of a rectangle,

$$\mathbb{P}(1.0 \leq X \leq 1.25) = (1.25-1.0) \times f(x) = (1.25-1.0) \times 2 = 0.5$$

14. From (Anderson et al. 2020) Chapter 6, #2

The random variable X is known to be uniformly distributed between 10 and 20 (in notation we write $X \sim \text{Unif}(10, 20)$)

(a) Show the graph of the probability density function.

(b) Compute $\mathbb{P}(X < 15)$ or Calculate $F(15)$.

(c) Compute $\mathbb{P}(12 \leq X \leq 18)$.

(d) Compute $\mathbb{E}(X)$.

(e) $\mathbb{V}(X)$.

Solution:

- a) This is a simple plot, we did a similar plot in the class

- b) Here $F(15) = \mathbb{P}(X < 15)$ means a cumulative probabilities upto 15. Let's derive the general formula to calculate the cumulative probability for a Uniform distribution with parameter a and b, The density function of the uniform is given by

$$f(x) = \frac{1}{b-a} \quad \text{for } a \leq x \leq b$$

Now we can get cumulative probability at c, with

$$F(c) = \mathbb{P}(X \leq c) = \int_a^c \frac{1}{b-a} dx = \frac{1}{b-a} \int_a^c dx = \frac{1}{b-a} [x]_a^c = \frac{c-a}{b-a}$$

Now we know that $a = 10$, $b = 20$, and $c = 15$, so we can calculate the cumulative probability at 15, finish it...Now to calculate probabilities for the uniform distribution we have this nice formula, but you cannot use this formula to calculate cumulative probabilities for any other distribution, this is only for the uniform distribution. For example, for the normal distribution we used the table in (Anderson et al. 2020).

- c) This is similar to a problem we solved in the class (and we have in the slides), we can solve it using a shortcut formula for the uniform, which is

$$\mathbb{P}(c \leq X \leq d) = (c - d) \times \frac{1}{b-a} = 1$$

In this case,

$$\mathbb{P}(12 \leq X \leq 18) = (18 - 12) \times \frac{1}{20 - 10} = 6 \times 0.1 = 0.6$$

Question is why does this formula work don't we need integral? The answer is if we do integral we will get the same answer,

$$\mathbb{P}(c \leq X \leq d) = \int_c^d \frac{1}{b-a} dx = \frac{1}{b-a} \int_c^d dx = \frac{1}{b-a} [x]_c^d = (d-c) \times \frac{1}{b-a}$$

if you draw it, you should see that this will be a rectangular area, and the area of a rectangle is the product of the length and width, which is $(d - c) \times \frac{1}{b-a}$.

Optional: We can also use cumulative probability to calculate this,

$$\mathbb{P}(12 \leq X \leq 18) = \mathbb{P}(X \leq 18) - \mathbb{P}(X \leq 12) = \frac{18-10}{20-10} - \frac{12-10}{20-10} = \frac{8}{10} - \frac{2}{10} = \frac{6}{10} = 0.6$$

- d and e. To solve d) and e) you can directly use the general formula for the Expectation and Variance of the Uniform distribution, so if $X \sim \mathcal{U}_{[a,b]}$, then

$$\mathbb{E}(X) = \frac{a+b}{2}$$
$$\mathbb{V}(X) = \frac{(b-a)^2}{12}$$

Now since we know $a = 10$ and $b = 20$, we can calculate $\mathbb{E}(X)$ and $\mathbb{V}(X)$, finish it.

Completely Optional (but good for understanding): You can also calculate the expectation and variance using the definition. In the sides we have a proof of expectation formula for the Uniform distribution. Let's see how can we derive the variance,

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] \\
&= \mathbb{E}\left[\left(X - \frac{a+b}{2}\right)^2\right] \\
&= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} dx \\
&= \frac{1}{b-a} \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx \\
&= \frac{1}{b-a} \int_a^b \left(x^2 - 2x\frac{a+b}{2} + \left(\frac{a+b}{2}\right)^2\right) dx \\
&= \frac{1}{b-a} \int_a^b x^2 dx - (a+b) \int_a^b x dx + \left(\frac{a+b}{2}\right)^2 \int_a^b dx \\
&= \frac{1}{b-a} \left[\frac{x^3}{3}\right]_a^b - (a+b) \left[\frac{x^2}{2}\right]_a^b + \left(\frac{a+b}{2}\right)^2 [x]_a^b \\
&= \ldots ... \\
&= \ldots ... \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

The calculation is very messy, don't worry it's not a math exam, so in the exam you won't have these calculations.... There is another way the calculation becomes a bit easier, note that now we know there is another way we can write variance, that is from question 5

$$\mathbb{V}(X) = \mathbb{E}\left[(X - \mathbb{E}\,(X))^2\right] = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

So we already know $\mathbb{E}(X) = \frac{a+b}{2}$, now we need to calculate $\mathbb{E}(X^2)$, this is a bit easier (apply LOTUS)

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_a^b x^2 \frac{1}{b-a} dx \\
&= \frac{1}{b-a} \int_a^b x^2 dx \\
&= \frac{1}{b-a} \left[\frac{x^3}{3}\right]_a^b \\
&= \frac{1}{3(b-a)} \left(b^3 - a^3\right) \\
&= \frac{1}{3(b-a)} (b-a) \left(b^2 + ab + a^2\right) \\
&= \frac{1}{3} \left(b^2 + ab + a^2\right)
\end{aligned}$$

So now

$$\mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{1}{3} \left(b^2 + ab + a^2\right) - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

So this means $\mathbb{V}(X) = \frac{(b-a)^2}{12}$, when we are calculating variance of the Uniform distribution, with parameters a and b

15. From (Anderson et al. 2020) Chapter 6, #9

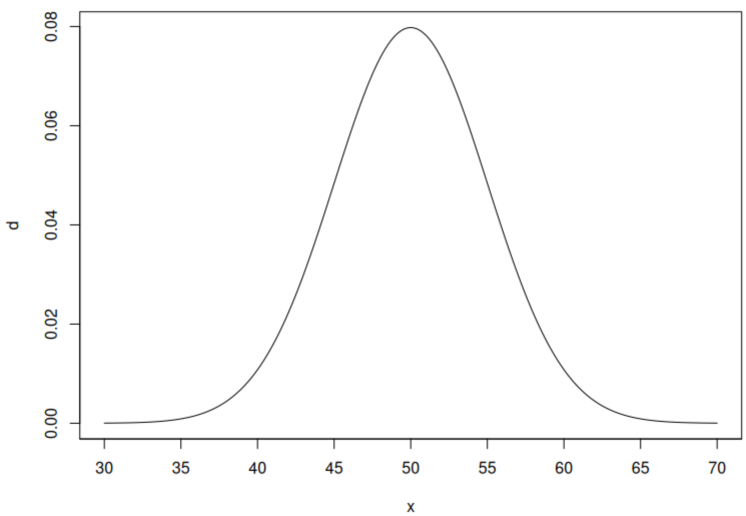
A random variable X is normally distributed with a mean of $\mu = 50$ and a standard deviation of $\sigma = 5$ (in notation we write $x \sim \mathcal{N}(50, 25)$)

(a) Sketch a normal curve for the probability density function (Using Figure 6.6 in (Anderson et al. 2020) as a guide). Label the horizontal axis with values of $35, 40, 45, 50, 55, 60$, and 65.

(b) What is the probability the random variable will assume a value between 45 and 55? This means $\mathbb{P}(45 < X < 55) = ?$

(c) What is the probability the random variable will assume a value between 40 and 60. This means $\mathbb{P}(40 < X < 60) = ?$

Solution: a). We did it in the class. This is a simple plot.



b) The solution is very easy,

$$
\begin{aligned}
\mathbb{P}(45 < X < 55) &= \mathbb{P}\left(\frac{45-50}{5} < \frac{X-\mu}{\sigma} < \frac{55-50}{5}\right) \\
&= \mathbb{P}\left(\frac{45-50}{5} < Z < \frac{55-50}{5}\right) \\
&= \mathbb{P}\left(-1 < Z < 1\right) \\
&= \mathbb{P}\left(Z < 1\right) - \mathbb{P}\left(Z < -1\right) \\
&= 0.8413 - 0.1587 = 0.6826
\end{aligned}
$$

In the second line we used Z transformation to bring everything into standard normal. The probabilities will be same, so we can do this. At the end we used the Z table in (Anderson et al. 2020) and found $\mathbb{P}\left(Z < 1\right) = 0.8413$ and $\mathbb{P}\left(Z < -1\right) = 0.1587$.

Interpretation: *So finally we calculated* $\mathbb{P}(45 < X < 55) = 0.6826$ *Question - What does* 0.6826 *mean? Or what is the interpretation of this number?*

Ans: We can comment on two ways,

- We can say that the probability of X taking values between 45 and 55 is 0.6826. Or there are 68.26% values of X that are within this range .
- We can also say that if we randomly pick 100 numbers from all possible values of X (which ranges from $-\infty$ to ∞), then almost 69 times we will get values between 45 and 65, the rest 31 times we won't.

Note that, the two comments are coming from classical definition and frequency definition of probability.

c) This is similar to b), you should try this! We will talk about this later.

16. From (Anderson et al. 2020) Chapter 6, #10

    Suppose we have a random variable which is distributed as a standard normal distribution (this means $Z \sim \mathcal{N}(0,1)$. Label the horizontal axis at values of $-3, -2, -1, 0, 1, 2$, and 3. Then use the table in (Anderson et al. 2020) to compute the following probabilities.

    (a) $\mathbb{P}(Z \leq 1.5)$
    (b) $\mathbb{P}(Z \leq 1)$
    (c) $\mathbb{P}(1 \leq Z \leq 1.5)$
    (d) $\mathbb{P}(0 < Z < 2.5)$

Solution: This is easy!

17. From (Anderson et al. 2020) Chapter 6, #12. Given that $Z \sim \mathcal{N}(0,1)$, compute the following probabilities.

    (a) $\mathbb{P}(0 \leq Z \leq .83)$
    (b) $\mathbb{P}(-1.57 \leq Z \leq 0)$
    (c) $\mathbb{P}(Z > .44)$
    (d) $\mathbb{P}(Z \geq -.23)$
    (e) $\mathbb{P}(Z < 1.20)$
    (f) $\mathbb{P}(Z \leq -.71)$

18. From (Anderson et al. 2020) Chapter 6, #15.Given that $Z \sim \mathcal{N}(0,1)$, find z for each situation.

    (a) The area to the left of z is .2119.
    (b) The area between $-z$ and z is .9030.
    (c) The area between $-z$ and z is .2052.
    (d) The area to the left of z is .9948.
    (e) The area to the right of z is .6915.

Solution:

- a) This is asking for the z value such that

$$\mathbb{P}(Z < z) = 0.2119$$

  Using the table we can directly calculate this, and this will be $z = -0.8$. This is the z value for which we have 21.19% cumulative probabilities. We have also learned that here -0.8 is called the 0.2119^{th} quantile or 21.19^{th} percentile.

- b) The question tells us that there is a z such that the area between $-z$ and z is .9030. So we have

$$\mathbb{P}(-z < Z < z) = 0.9030$$

This is the probability in center, the probabilities in two tails would be $1 - .9030 = 0.097$ This means the tail probabilities together is 0.097. Now since the distribution is symmetric around 0. To get the probability in the left tail, we can simply divide this number by 2, so we have

$$\mathbb{P}(Z < z^*) = 0.097/2 = 0.0485$$

here z^* is the z value for which we have 0.0485 cumulative probabilities (this means probabilities on the left). If you see the table then this value is -1.66. So now we know that

$$\mathbb{P}(-1.66 < Z < 1.66) = 0.9030$$

So here

  - $-z = -1.66$ and $z = 1.66$

c) Same as b)

d) We need to find z such that

$$\mathbb{P}(Z < z) = .9948$$

Finish this....

e) We need to find z such that

$$\mathbb{P}(Z > z) = .6915$$

but since we know

$$\mathbb{P}(Z > z) = 1 - \mathbb{P}(Z < z)$$

so we can write

$$\begin{aligned} 1 - \mathbb{P}(Z < z) &= .6915 \\ \mathbb{P}(Z < z) &= 1 - .6915 = 0.3085 \end{aligned}$$

Now you should be able to figure out z, and from the table we get $z = -.5$.

19. (**SKIP THIS FOR EXAM**) From (Anderson et al. 2020) Chapter 6, #33.

    Consider a random variable following exponential probability density function.

    $$f(x) = \frac{1}{3}e^{-x/3} \quad \text{for } x \geq 0$$

    (a) Write the formula for $P(x \leq x_0)$.
    (b) Find $\mathbb{P}(X \leq 2)$.
    (c) Find $\mathbb{P}(X \geq 3)$.
    (d) Find $\mathbb{P}(X \leq 5)$.
    (e) Find $\mathbb{P}(2 \leq X \leq 5)$.

20. From (Anderson et al. 2020) Chapter 6, #3.

    Cincinnati to Tampa Flight Time. Delta Airlines quotes a flight time of 2 hours, 5 minutes for its flights from Cincinnati to Tampa. Suppose we believe that actual flight times are uniformly distributed between 2 hours and 2 hours, 20 minutes.

    (a) Show the graph of the probability density function for flight time.
    (b) What is the probability that the flight will be no more than 5 minutes late?
    (c) What is the probability that the flight will be more than 10 minutes late?
    (d) What is the expected flight time?

Solution:

- a. First of all note that here $a = 120$ and $b = 140$. Now we can use the formula for the Uniform distribution to calculate the density function,

$$f(x) = \frac{1}{b-a} = \frac{1}{140-120} = 0.05$$

- No more than 5 minutes late means it must be less than $120 + 5 = 125$ mins, so this means we need to calculate

$$\mathbb{P}(X \leq 125) = 0.05 \times (125 - 120) = 0.05 \times 5 = 0.25$$

- More than 10 min late means $\mathbb{P}(X \geq 135)$, so in this case

$$\mathbb{P}(X \geq 135) = 0.05 \times (140 - 135) = 0.05 \times 5 = 0.25$$

- We can use the formula for the expectation of the Uniform distribution,

$$\mathbb{E}(X) = \frac{a+b}{2} = \frac{120+140}{2} = 130$$

21. From (Anderson et al. 2020) Chapter 6, #7.

Bidding on Land. Suppose we are interested in bidding on a piece of land and we know one other bidder is interested. The seller announced that the highest bid in excess of \$10,000 will be accepted. Assume that the competitor's bid X is a random variable that is uniformly distributed between \$10,000 and \$15,000.

(a) Suppose you bid \$12,000. What is the probability that your bid will be accepted?
(b) Suppose you bid \$14,000. What is the probability that your bid will be accepted?
(c) What amount should you bid to maximize the probability that you get the property?

Solution:

Here your competitor's bid X is uniformly distributed between \$10,000 and \$15,000, so we can write

$$X \sim \textbf{Unif}(10000, 15000)$$

The density function in this case is,

$$f(x) = \frac{1}{15000 - 10000} = \frac{1}{5000} \quad \text{for } 10000 \leq x \leq 15000$$

(a) Recall here X is the competitor's bid. It's a bidding problem, in this case you will win if your competitor's bid is less than your bid. So if you bid \$12,000, then you will win if $X < 12000$. So we need to calculate

$$\mathbb{P}(X < 12000) = \frac{1}{5000} \times (12000 - 10000) = \frac{1}{5000} \times 2000 = 0.4$$

So the probability of your bid being accepted is 0.4 or 40%.

(b) Similar to the last one, in this case your competitor needs to bid less than 14,000, so

$$\mathbb{P}(X < 14000) = \frac{1}{5000} \times (14000 - 10000) = \frac{1}{5000} \times 4000 = 0.8$$

So the probability of your bid being accepted is 0.8 or 80%.

(c) Obviously if your competitor's bidding range is between 10,000 and 15,000, then you should bid more than 15,000 to maximize the probability of winning. So the maximum probability of winning is 1 or 100% if you bid more than 15,000, since

$$\mathbb{P}(X < 15000) = \frac{1}{5000} \times (15000 - 10000) = \frac{1}{5000} \times 5000 = 1$$

22. From (Anderson et al. 2020) Chapter 6, #17.

Height of Dutch Men. Males in the Netherlands are the tallest, on average, in the world with an average height of 183 centimeters (cm) (BBC News website). Assume that the height of men in the Netherlands is normally distributed with a mean of 183 cm and standard deviation of 10.5 cm.

(a) What is the probability that a Dutch male is shorter than 175 cm ?

(b) What is the probability that a Dutch male is taller than 195 cm ?

(c) What is the probability that a Dutch male is between 173 and 193 cm ?

(d) Out of a random sample of 1000 Dutch men, how many would we expect to be taller than 190 cm ?

Solution:

In this case we assume the height of the Dutch men X is normally distributed with mean $\mu = 183$ cm and standard deviation $\sigma = 10.5$ cm, so this means

$$X \sim \mathcal{N}(183, 10.5^2)$$

(a) Here we need to calculate

$$\begin{aligned}\mathbb{P}(X < 175) &= \mathbb{P}\left(\frac{X-183}{10.5} < \frac{175-183}{10.5}\right) \\ &= \mathbb{P}\left(Z < -0.76\right) = 0.2236 \text{ (We get this from the table)}\end{aligned}$$

(b) Here we need to calculate

$$\begin{aligned}\mathbb{P}(X > 195) &= \mathbb{P}\left(\frac{X-183}{10.5} > \frac{195-183}{10.5}\right) \\ &= \mathbb{P}\left(Z > 1.14\right) = 1 - \mathbb{P}\left(Z < 1.14\right) = 1 - 0.8729 = 0.1271\end{aligned}$$

(c) Here we need to calculate

$$\begin{aligned}\mathbb{P}(173 < X < 193) &= \mathbb{P}(\frac{173-183}{10.5} < \frac{X-\mu}{\sigma} < \frac{193-183}{10.5}) \\ &= \mathbb{P}(-0.95 < Z < 0.95) \\ &= \mathbb{P}(Z < 0.95) - \mathbb{P}(Z < -0.95) \\ &= 0.8289 - 0.1711 = 0.6578\end{aligned}$$

(d) For this problem first find the probability,

$$\begin{aligned}\mathbb{P}(X > 190) &= \mathbb{P}\left(\frac{X-183}{10.5} > \frac{190-183}{10.5}\right)\\ &= \mathbb{P}\left(Z > 0.67\right) = 1 - \mathbb{P}\left(Z < 0.67\right) = 1 - 0.7486 = 0.2514 = 25.14\%\end{aligned}$$

So this means out of 100 there are around 25 people who are taller than 190 cm. So out of 1000 people we expect around 250 people who are taller than 190 cm.

23. From (Anderson et al. 2020) Chapter 6, #20.

Gasoline Prices. Suppose that the average price for a gallon of gasoline in the United States is \$3.73 and in Russia is \$3.40. Assume these averages are the population means in the two countries and that the probability distributions are normally distributed with a standard deviation of \$.25 in the United States and a standard deviation of \$.20 in Russia.

(a) What is the probability that a randomly selected gas station in the United States charges less than \$3.50 per gallon?

(b) What percentage of the gas stations in Russia charge less than \$3.50 per gallon?

(c) What is the probability that a randomly selected gas station in Russia charged more than the mean price in the United States?

Solution:

So here we have two random variables, X_1 is the gasoline price of US and X_2 is the gasoline price of Russia. We know that both of these random variables are normally distributed,

In the case of US we know the mean of US gasoline price is $\mu_1 = 3.73$ and standard deviation is $\sigma_1 = 0.25$. So we can write

$$X_1 \sim \mathcal{N}(3.73, 0.25^2)$$

and for Russia we know the mean is $\mu_2 = 3.40$ and standard deviation is $\sigma_2 = 0.20$. So we can write

$$X_2 \sim \mathcal{N}(3.40, 0.20^2)$$

(a) This is for US

We need to calculate

$$\begin{aligned}\mathbb{P}(X_1 < 3.50) &= \mathbb{P}\left(\frac{X_1-3.73}{0.25} < \frac{3.50-3.73}{0.25}\right)\\ &= \mathbb{P}\left(Z < -0.92\right) = 0.1788\end{aligned}$$

So the probability that a randomly selected gas station in the United States charges less than \$3.50 per gallon is 0.1788 or 17.88%.

(b) This is for Russia

We need to calculate

$$\begin{aligned}\mathbb{P}(X_2 < 3.50) &= \mathbb{P}\left(\frac{X_2 - 3.40}{0.20} < \frac{3.50 - 3.40}{0.20}\right) \\ &= \mathbb{P}\left(Z < 0.50\right) = 0.6915\end{aligned}$$

So the percentage of the gas stations in Russia charge less than \$3.50 per gallon is 0.6915 or 69.15%.

(c) Now the last question asks for Russia, when it charges more than US mean or more than $\mu_1 = \$3.73$. So we need to calculate

$$\begin{aligned}\mathbb{P}(X_2 > \mu_1) = \mathbb{P}(X_2 > 3.73) = \mathbb{P}\left(\frac{X_2 - 3.40}{0.20} > \frac{3.73 - 3.40}{0.20}\right) = \\ = \mathbb{P}\left(Z > 1.65\right) = 1 - \mathbb{P}\left(Z < 1.65\right) = 1 - 0.9505 = 0.0495\end{aligned}$$

24. From (Anderson et al. 2020) Chapter 6, #19.

Automobile Repair Costs. Automobile repair costs continue to rise with an average 2015 cost of \$367 per repair (U.S. News & World Report website). Assume that the cost for an automobile repair is normally distributed with a standard deviation of \$88. Answer the following questions about the cost of automobile repairs.

(a) What is the probability that the cost will be more than \$450 ?

(b) What is the probability that the cost will be less than \$250 ?

(c) What is the probability that the cost will be between \$250 and \$450 ?

(d) If the cost for your car repair is in the lower 5% of automobile repair charges, what is your cost?

Solution:

Here automobile repair costs are normally distributed with a mean of $\mu = 367$ and standard deviation of $\sigma = 88$. We can write this as $X \sim \mathcal{N}(367, 88^2)$, where X is a random variable which represents automobile costs.

- a) we need find $\mathbb{P}(X > 450)$. Use Z transformation and standard normal table
- b) we need find $\mathbb{P}(X < 250)$. Again use Z transformation and standard normal table.

- c) we need find $\mathbb{P}(250 < X < 450)$. Again use Z transformation and standard normal table.
- d) we need find x such that $\mathbb{P}(X < x) = 0.05$. But we cannot directly find this x, because we don't have the table for $\mathcal{N}(367, 88^2)$, we only have table for $\mathcal{N}(0, 1)$. First, we need to find z such that $\mathbb{P}(Z < z) = 0.05$, then we can use Z transformation to find x.
  - step 1: Let's find z such that $\mathbb{P}(Z < z) = 0.05$. From the table we see that $z = -1.645$.
  - step 2: Now we can use Z transformation to find x,

  $$x = \mu + \sigma z = 367 + 88 \times (-1.645) = 225.82$$

  Why we can do this? We can do this because of the relation between standard normal and any other normal distribution. The probabilities (or area) are same, this means

  $$\underbrace{\mathbb{P}(X < 225.82)}_{\text{this is probability under } \mathcal{N}(367,88^2)} = \underbrace{\mathbb{P}(Z < -1.645)}_{\text{this is probability under } \mathcal{N}(0,1)}$$

So if I know the cost for my car repair is in the lower 5% of automobile repair charges, then I know my cost would be less than \$225.82.

25. (**SKIP THIS FOR EXAM**) From (Anderson et al. 2020) Chapter 6, #34.

    Phone Battery Life. Battery life between charges for a certain mobile phone is 20 hours when the primary use is talk time, and drops to 7 hours when the phone is primarily used for Internet applications over a cellular network. Assume that the battery life in both cases follows an exponential distribution.

    (a) Show the probability density function for battery life for this phone when its primary use is talk time.

    (b) What is the probability that the battery charge for a randomly selected phone will last no more than 15 hours when its primary use is talk time?

    (c) What is the probability that the battery charge for a randomly selected phone will last more than 20 hours when its primary use is talk time?

    (d) What is the probability that the battery charge for a randomly selected phone will last no more than 5 hours when its primary use is Internet applications?

26. (**SKIP THIS FOR EXAM**) From (Anderson et al. 2020) Chapter 6, #38.

    Boston 911 Calls. The Boston Fire Department receives 911 calls at a mean rate of 1.6 calls per hour (Mass.gov website). Suppose the number of calls per hour follows a Poisson probability distribution.

(a) What is the mean time between 911 calls to the Boston Fire Department in minutes?

(b) Using the mean in part (a), show the probability density function for the time between 911 calls in minutes.

(c) What is the probability that there will be less than one hour between 911 calls?

(d) What is the probability that there will be 30 minutes or more between 911 calls?

(e) What is the probability that there will be more than 5 minutes, but less than 20 minutes between 911 calls?

Remarks: All problems are taken from or modified from (Anderson et al. 2020). If possible you should do more problems from there.

References

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., Cochran, J. J., Fry, M. J. & Ohlmann, J. W. (2020), *Statistics for Business & Economics*, 14th edn, Cengage, Boston, MA.