

STAT 2290 Homework 4 (corrected)

Hand in on April 1 or 2 with your quiz.

Read these instructions:

- Show your work. Answers without proper justifications receive no credit.
- The only integration techniques we will use are: power rule, substitution rule, exponential functions.

Types of random variables

Problem 1.

Let X be the random variable that records the number of times that "1" comes up when a fair dice is rolled 4 times.

(a) State the values of $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X = 4)$.

(b) Draw the histogram of the PMF of X .

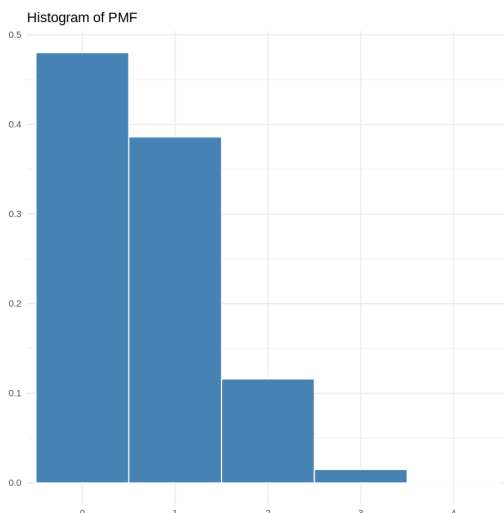
(c) Draw the histogram of the CDF of X .

Solution.

This X is a binomial random variable, which we can think of as the result of flipping a coin 4 times, with probability $p = \frac{1}{6}$ of heads (corresponding to rolling a "1"):

- $P(X = 0) = \binom{4}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^4 \approx 0.48$
- $P(X = 1) = \binom{4}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^3 \approx 0.386$
- $P(X = 2) = \binom{4}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 \approx 0.116$
- $P(X = 3) = \binom{4}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^1 \approx 0.015$
- $P(X = 4) = \binom{4}{4} \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^0 \approx 0.00077$

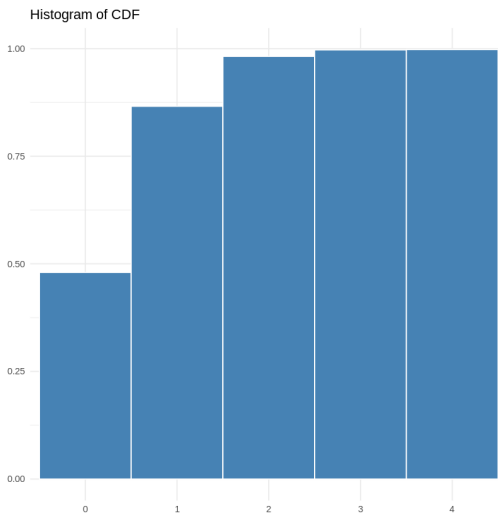
Therefore the PMF looks like this as a histogram:



The histogram of the CDF of X therefore has bars of heights:

- 0.48,
- $0.48+0.386$,
- $0.48+0.386+0.116$,

- $0.48+0.386+0.116+0.015$,
- $0.48+0.386+0.116+0.015+0.00077 = 1$



Problem 2.

Each of the 10 questions on an exam has multiple choice options A, B, C, D, and E. Let X be the random variable that records the number of correct answers that result when you randomly guess on each of the 10 questions.

- Compute $P(X \geq 9)$.
- Compute $F(1)$ if $F(x)$ is the CDF of X .

Solution.

This X is a binomial random variable with 10 flips of the coin and probability $p = 1/5$ of heads (aka 10 trials with probability $p = 1/5$ of success).

Since there are only n trials, for (a):

$$P(X \geq 9) = P(X = 9) + P(X = 10) = \binom{10}{9} \left(\frac{1}{5}\right)^9 \left(\frac{4}{5}\right)^1 + \binom{10}{10} \left(\frac{1}{5}\right)^{10} \left(\frac{4}{5}\right)^0$$

For (b):

$$F(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = \binom{10}{0} \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10} + \binom{10}{1} \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^9$$

Problem 3.

You roll a fair dice many times, and you only stop rolling once you get a non-"1" from a dice roll. Let Z be the random variable that records the number of "1"s you get before you finally get a non-"1" dice roll.

- Compute $f(2)$ if $f(z)$ is the PMF of Z .
- Compute $F(2)$ if $F(z)$ is the CDF of Z .
- Find the probability that it takes at least 3 rolls before you get a non-"1".
(An earlier version of (c) Find the probability that it takes at least 3 rolls before you get a "1".)

Solution.

$Z \sim \text{Geometric}(p = 5/6)$ is a geometric random variable with probability $p = 5/6$ of success.

- $f(2) = P(Z = 2) = (1/6)^2(5/6)$.
- $F(2) = P(Z \leq 2) = P(Z = 0) + P(Z = 1) + P(Z = 2) = (1/6)^0(5/6) + (1/6)^1(5/6) + (1/6)^2(5/6)$
- The words "at least" with many cases to consider suggest counting by complement:

$$P(Z \geq 3) = 1 - P(Z \leq 2) = 1 - [(1/6)^0(5/6) + (1/6)^1(5/6) + (1/6)^2(5/6)]$$

Problem 4.

Consider the Pareto random variable X whose PDF has the form $f(x) = c/x^2$ for $x \geq 1$ and $f(x) = 0$ otherwise.

- Find c .

(b) Find the CDF of X .

(c) Find the median of X .

Solution.

(a):

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_1^{\infty} cx^{-2} dx = [-cx^{-1}]_1^{\infty} = 0 - (-c) = c \Rightarrow \boxed{c=1}$$

(b): There are two cases: if $x \leq 1$, then $F(x) = 0$; and if $x \geq 1$, then

$$F(x) = \int_{-\infty}^x f(t) dt = \int_1^x t^{-2} dt = [-t^{-1}]_1^x = -\frac{1}{x} - \left(-\frac{1}{1}\right) = 1 - \frac{1}{x}.$$

(c): The median occurs at $0.5 = F(x) = 1 - \frac{1}{x}$ so $x = 2$.

Problem 5.

The time to failure (in hours) of a PC component has PDF $f(x) = \begin{cases} \frac{1}{1000} \exp(-x/1000) & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$

(a) Find the CDF $F(x)$.

(b) Find the probability that the component lasts more than 1000 hours.

(c) Find the probability that the component fails before 2000 hours.

Note: $\frac{1}{1000} \exp(-x/1000) = \frac{1}{1000} e^{-x/1000}$

Solution.

(a): There are two cases: if $x \leq 0$, then $F(x) = 0$; and if $x \geq 0$, then

$$F(x) = \int_{-\infty}^x f(t) dt = \int_0^x \frac{1}{1000} e^{-t/1000} dt = [-e^{-t/1000}]_0^x = 1 - e^{-x/1000}$$

(b): Let X be the random variable in question.

$$P(X > 1000) = 1 - P(X \leq 1000) = 1 - F(1000) = 1 - [1 - e^{-1000/1000}] = e^{-1} \approx 36.8\%.$$

(c):

$$P(X \leq 2000) = F(2000) = 1 - e^{-2000/1000} = 1 - e^{-2} \approx 68\%.$$

Problem 6.

The proportion of people who respond to a mail-order solicitation is a continuous random variable X that has the probability density function $f(x) = 2(x+2)/5$ for $0 \leq x \leq 1$ and $f(x) = 0$ elsewhere.

(a) Show that $P(0 \leq x \leq 1) = 1$.

(b) Find the probability that more than 1/4 but fewer than 3/4 of the people contacted will respond to this type of solicitation.

Solution.

(a):

$$\int_0^1 \frac{2(x+2)}{5} dt = \int_0^1 \frac{2x}{5} + \frac{4}{5} dt = \left[\frac{x^2}{5} + \frac{4x}{5} \right]_0^1 = \left[\frac{1^2}{5} + \frac{4 \cdot 1}{5} \right] - \left[\frac{0^2}{5} + \frac{4 \cdot 0}{5} \right] = 1$$

(b):

$$P(1/4 \leq X \leq 3/4) = \int_{1/4}^{3/4} \frac{2x}{5} + \frac{4}{5} dt = \left[\frac{x^2}{5} + \frac{4x}{5} \right]_{1/4}^{3/4} = 0.5$$

Expected value and variance

Problem 7.

The probability distribution of the discrete random variable, X , is

$$f(x) = \binom{5}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{5-x}, \quad x = 0, 1, 2, 3, 4, 5.$$

(a) Find the mean of X .

(b) Find the standard deviation of X .

(Hint: Is this the PMF of a special type of RV?)

Solution.

This is the PMF of a binomial random variable X with $n = 5$ flips of the coin and probability $p = 1/4$ of success. Therefore by the expected value and variance formulas for the binomial random variable:

(a) $E(X) = np = 5 \cdot 1/4 = 5/4$

(b) $\sigma_X = \sqrt{np(1-p)} = \sqrt{5 \cdot (1/4) \cdot (3/4)}$.

Problem 8.

You repeatedly draw cards from a standard deck of 52 cards.

(Imagine you are playing blackjack and you repeatedly "hit".)

Let X be the random variable that records the number of non-heart cards you receive before finally receiving a heart card.

(a) Find Approximate $E(X)$ using a geometric random variable.

(b) Find Approximate σ_X using a geometric random variable.

(c) Explain why the geometric random variable only gives an approximate answer to (a) and (b) above.

Solution.

We approximate X by the geometric random variable with probability p of getting a heart card out of 52 cards, or $p = 1/4$.

(a) $E(X) \approx 1/p - 1 = 1/(1/4) - 1 = 4 - 1 = 3$

(b) $\sigma_X^2 \approx (1-p)/p^2 = (3/4)/(1/4)^2 = 12$ so $\sigma_X = \sqrt{12}$ (Corrected answer on 2026 April 9th)

(c) Two reasons:

- First, the geometric random variable requires possibly infinite number of tries before a first success, but here we only have 52 cards before we run out of tries (since we sample without replacement).
- Second, the geometric random variable assumes sampling without replacement, which gives different probabilities from sampling with replacement. (The lesson on sampling distributions, we described when sampling with replacement and sampling without replacement produce approximately similar probabilities.)

Problem 9.

The proportion of people who respond to a mail-order solicitation is a continuous random variable X that has the probability density function $f(x) = 2(x+2)/5$ for $0 \leq x \leq 1$ and $f(x) = 0$ elsewhere.

(a) Find $E(X)$.

(b) Find σ_X .

Solution.

(a)

$$E(X) = \int_0^1 x \left(\frac{2x}{5} + \frac{4}{5} \right) dx = \int_0^1 \frac{2x^2}{5} + \frac{4x}{5} dx = \left[\frac{2x^3}{15} + \frac{2x^2}{5} \right]_0^1 = \frac{8}{15}$$

(b)

$$E(X^2) = \int_0^1 x^2 \left(\frac{2x}{5} + \frac{4}{5} \right) dt = \int_0^1 \frac{2x^3}{5} + \frac{4x^2}{5} dt = \left[\frac{x^4}{10} + \frac{4x^3}{15} \right]_0^1 = \frac{11}{30}$$

So $\sigma_X^2 = E(X^2) - E(X)^2 = (11/30) - (8/15)^2 \approx 0.082$

And so $\sigma_X \approx \sqrt{0.082}$.

Problem 10.

The random variable X measures the heights in inches of US men between ages 20-29. We know that $E(X) = 70$ inches and $\sigma_X = 3$ inches. Re-express $E(X)$ and σ_X in terms of centimeters.

Solution.

Let Y be the same random variable but in centimeters.

Since 1 inch is approximately 2.5 cm, $Y = 2.5X$ and so

$$E(Y) = 2.5 \cdot E(X) = 2.5 \cdot 70 = 175 \text{ cm}$$

$$\sigma_Y = 2.5 \cdot \sigma_X = 2.5 \cdot 3 = 7.5 \text{ cm}$$

Problem 11.

The random variable Y measures the average daily temperature of each day in Glassboro in 2025. We know that $E(Y) = 55^\circ F$ and $\sigma_Y = 18^\circ F$. Re-express these in terms of Celsius.

Solution.

Let Z be the random variable Y expressed in Celsius.

Celsius is related to Fahrenheit via the formula $^\circ C = \frac{5}{9}(F - 32)$ so

$$E(Z) = E\left(\frac{5}{9}(Y - 32)\right) = \frac{5}{9}(E(Y) - 32) = \frac{5}{9}(55 - 32) = \frac{5 \cdot 23}{9} \approx 12.7^\circ C$$

$$\sigma_Z = \frac{5}{9}\sigma_Y = \frac{5}{9} \cdot 18 = 10^\circ C$$