"You Do It" Solution

Lecture 5: Probability Distributions

Prepared as a classroom-ready detailed solution set for ECOM5005

Exercise 1: Guessing on a Multiple-Choice Exam

Question

A student decides to guess on the multiple-choice section of the ECOM5005 final exam. The section contains **50 questions**, and each question has **5 possible answers**.

(a) Find the expected number of correct responses.

(b) Find the standard deviation of the number of correct responses (keep 4 decimal places).

Detailed Solution

Step 1: Identify the appropriate probability distribution.
Let X be the number of questions answered correctly. Because there is a fixed number of trials, each question has only two outcomes for our purpose (correct or incorrect), the probability of a correct answer is constant, and the guesses are treated as independent, X follows a binomial distribution:

$$X \sim \text{Bin}(n = 50, p = 0.2).$$

Since there are 5 possible answers and only one is correct,

$$p = P(\text{correct}) = \frac{1}{5} = 0.2, \qquad 1 - p = 0.8.$$

Step 2: Calculate the expected number of correct responses.
For a binomial random variable,

$$\mu = E(X) = np.$$

Therefore,

$$E(X) = 50(0.2) = \boxed{10}.$$

So, if many students guessed on 50 such questions, the long-run average number of correct answers would be 10.

Step 3: Calculate the variance.
For a binomial random variable,

$$\sigma^2 = np(1 - p).$$

Thus,

$$\sigma^2 = 50(0.2)(0.8) = 8.$$

Step 4: Calculate the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{8} = 2.828427\ldots$$

To four decimal places,

$$\boxed{\sigma = 2.8284}.$$

Exercise 1: Interpretation and Final Answer

Why the Binomial Model Applies

The random variable counts the number of **correct answers** out of a fixed total of 50 questions. Each guessed question is treated as a Bernoulli trial with the same success probability $p = 0.2$. Therefore, the binomial mean and standard-deviation formulas from Lecture 5 apply directly.

Expected Value versus a Guaranteed Outcome

The result

$$E(X) = 10$$

does **not** mean that a student who guesses must get exactly 10 questions correct. The expected value is a **long-run average**. One student might get 6 correct, another 12, another 15, and so on; across many repetitions, the average number correct would approach 10.

Interpreting the Standard Deviation

The standard deviation

$$\sigma = 2.8284$$

measures the typical spread of the number of correct answers around the mean of 10. A result several standard deviations away from 10 would be relatively unusual under pure random guessing.

Answer Summary

(a) Expected number of correct responses:

$$\boxed{E(X) = np = 50(0.2) = 10.}$$

(b) Standard deviation:

$$\boxed{\sigma = \sqrt{50(0.2)(0.8)} = \sqrt{8} = 2.8284.}$$

Final Check

For a binomial distribution, remember the three related formulas:

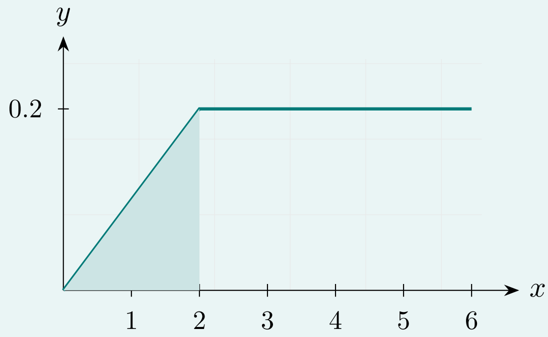
$$\mu = np, \qquad \sigma^2 = np(1-p), \qquad \sigma = \sqrt{np(1-p)}.$$

A common error is to report the variance 8 when the question asks for the **standard deviation**. The standard deviation is the square root of the variance, $\sqrt{8} = 2.8284$.

Exercise 2: Probability from a Density Curve

Question

Consider the density curve below. Find the probability that x is less than 2.



Detailed Solution

Step 1: Recall how probabilities are obtained from a continuous density curve.
For a continuous random variable, probability is represented by **area under the density curve**. Therefore,

$$P(X < 2) = \text{area under the curve from } x = 0 \text{ to } x = 2.$$

Step 2: Identify the geometric shape of the required region.
From the graph, the density increases linearly from $(0, 0)$ to $(2, 0.2)$. Hence the region under the curve for $0 < x < 2$ is a triangle with

$$\text{base} = 2, \qquad \text{height} = 0.2.$$

Step 3: Calculate the area of the triangle.
Using

$$\text{Area of a triangle} = \frac{1}{2}(\text{base})(\text{height}),$$

we obtain

$$P(X < 2) = \frac{1}{2}(2)(0.2) = 0.2.$$

Therefore,

$$\boxed{P(X < 2) = 0.20}.$$

Step 4: Check that the entire density integrates to 1.
The area from 0 to 2 is 0.2. From 2 to 6, the graph is a rectangle of height 0.2 and width 4, so its area is

$$4(0.2) = 0.8.$$

Thus the total area is

$$0.2 + 0.8 = 1,$$

confirming that the graph is a valid probability density curve.

Exercise 2: Interpretation and Final Answer

Probability Is Area, Not Height

At $x = 2$, the density height is 0.2, but this is **not** itself the probability that $X < 2$. For a continuous random variable, probabilities are obtained by calculating areas over intervals. In this problem the numerical answer also happens to be 0.2, but it comes from the triangular area, not simply from reading the height of the curve.

Geometric Calculation

The required region is the triangle to the left of $x = 2$:

$$P(X < 2) = \frac{1}{2} \times \underbrace{2}_{\text{base}} \times \underbrace{0.2}_{\text{height}} = 0.20.$$

So **20% of the total probability area** lies below $x = 2$.

Equivalent Calculus Check

The rising part of the density curve has slope

$$\frac{0.2 - 0}{2 - 0} = 0.1,$$

so for $0 \le x \le 2$ the density can be written as $f(x) = 0.1x$. Therefore,

$$P(X < 2) = \int_0^2 0.1x\,dx = 0.05x^2\Big|_0^2 = 0.20.$$

This gives the same result as the geometric area method.

Answer Summary

Final answer:

$$\boxed{P(X < 2) = 0.20 = 20\%.}$$

Reason: The probability equals the area of the triangular region under the density curve from $x = 0$ to $x = 2$.

Final Check

For continuous distributions, remember:

- probability is **area under the density curve**;
- the total area under a valid density curve equals 1;
- the probability of any exact single value is 0, so $P(X < 2) = P(X \le 2)$ for a continuous random variable.