Expectations and Variance Exercises (Pre-Reading for Data Science MSc)

1. Define expectation for a discrete random variable. Give an example with a fair die.

Expectation: most likely outcome (mean)

$E[X] = \sum_x x P(X=x)$

Eg.: expectation for a fair die = 3.5

2. Compute $E[X]$ for a fair coin toss where $X = 1$ for heads, 0 for tails.

$$E[X] = \sum_x x P(X=x) = 1 \times 0.5 + 0 \times 0.5 = 0.5$$

3. Compute $E[X]$ for rolling a fair six-sided die.

$$E[X] = \sum_x x P(X=x) = 1\left(\tfrac{1}{6}\right) + 2\left(\tfrac{1}{6}\right) \dots 6\tfrac{1}{6} = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$$

4. Define expectation for a continuous random variable with PDF $f(x)$.

$$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$$

5. For $X \sim U(0,1)$, compute $E[X]$.

↳ X has uniform dist. on interval [0,1]

$$\int_0^1 x f(x)\, dx = \int_0^1 x\, dx = \left[\frac{x^2}{2}\right]_0^1 = 0.5$$

6. State and explain the linearity of expectation property.

$E[aX + bY] = aE[X] + bE[Y]$

Expectation distributes across sums / scalars

7. Let X be die roll ($E[X] = 3.5$), Y be coin toss ($E[Y] = 0.5$). Compute $E[2X + 3Y]$.

$E[2X+3Y] = 2 \times 3.5 + 3 \times 0.5 = 8.5$

8. Define variance and explain its interpretation.

Variance (σ^2) = squared deviation from mean
↳ measures spread of a distribution

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

$\text{Var}[X] = \sigma^2$

$\sigma = X - E[X] \rightarrow \sigma^2 = E[(X - E[X])^2]$

$= E[X^2] - (E[X])^2$

10. Compute $\text{Var}(X)$ for a fair die.

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

$E[X] = 3.5 \qquad (E[X])^2 = 12.25$

$E[X^2] = \dfrac{1^2 + 2^2 + \ldots + 6^2}{6} \approx 15.17$

$\therefore \text{Var}(X) = 2.92$

11. State how variance changes under scaling and shifting: $\mathrm{Var}(aX + b)$.

$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$

- Shifting $\Rightarrow$ unaffected
- Scaling $\rightarrow$ multiplied by squared scale factor

12. If $\mathrm{Var}(X) = 4$, compute $\mathrm{Var}(3X + 2)$.

$3^2 \times 4 = 36$

13. State formula for variance of sums of independent variables.

$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

↳ if X, Y are independent

14. Compute $\mathrm{Var}(X + Y)$ where X, Y are independent dice rolls.

$\mathrm{Var}(X) = E[(\cancel{E}X - E[X])^2]$

For X: $E[X^2] - (E[X])^2$

$= 15.17 - 12.25 = 2.92$

$2.92 \cdot 2 = 5.84$

15. Define standard deviation and explain why it is more interpretable than variance.

$SD = \sqrt{\mathrm{Var}} = \sigma$

SD is in original units, interpretable as average distance from mean

16. Define covariance. Interpret positive, negative, and zero covariance.

Positive → Move in same direction
Negative → Move in opposite directions
Zero → No linear relation

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$

17. If X and Y are independent, what is $\text{Cov}(X, Y)$?

0

18. Define correlation $\rho(X, Y)$ and state its range of values.

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}, \text{ always in } [-1, 1]$$

19. Interpret correlation values: $\rho = +1$, $\rho = -1$, $\rho = 0$.

$\rho = +1$ → Perfectly positive line
$\rho = -1$ → Perfectly negative line
$\rho = 0$ → No linear relation

20. In a dataset, feature A has mean 10 and SD 2, feature B has mean 20 and SD 5, covariance $\text{Cov}(A, B) = 5$. Compute $\rho(A, B)$.

$$\rho(A, B) = \frac{5}{10} = 0.5$$