Main class 4: Expectation, Variance, Joint Random Variables

Izzy Grosof

September 24th, 2025

1 Summary

- **Definition 1.** The expected value $\mathbb{E}[X]$ (mean) of a r.v. X is
  - Discrete r.v.: $\mathbb{E}[X] = \sum_x x\mathbb{P}\{X = x\}$
  - Continuous r.v.: $\mathbb{E}[X] = \int_x x f_X(x)dx$

  If neither, other ways to define.
- Linearity:
$$\mathbb{E}[aX] = a\mathbb{E}[X]$$
$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$
- **Definition 2.** The variance $\mathrm{Var}[X]$ of a r.v. X is defined to be $\mathrm{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[Y^2]$
$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$
- Scaling:
$$\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$$
$$\mathrm{Var}[X + b] = \mathrm{Var}[X]$$
- For pairs of random variables X, Y, we define joint CDFs, joint pmf (if both discrete), and joint densities (if both continuous):
  - Joint CDF: $F_{X,Y}(x, y) = \mathbb{P}\{X \le x, Y \le y\}$
  - If both discrete: Joint pmf: $\mathbb{P}\{X = x, Y = y\}$.
  - If both continuous: Joint density: $f_{X,Y}(x, y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}F_{X,Y}(x, y)$

  "∂" is the partial derivative symbol, and it's pronounced "Partial".
- A "marginal" CDF, pmf, or density is the single-random-variable version.
- We define random variables X and Y to be independent, $X \perp Y$, if
$$F_{X,Y}(x, y) = F_X(x)F_Y(y)$$
$$\mathbb{P}\{X \le x, Y \le y\} = \mathbb{P}\{X \le x\}\mathbb{P}\{Y \le y\}$$
- For any functions g, h, if $X \perp Y$, then
$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$$
- Covariance: $Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

2 Clarifications & corrections from last class

- Forgot to post textbook chapters and course notes until later on Tuesday. Sorry, will be sooner.
- For continuous random variables, the CDF is continuous - it doesn't need to be differentiable. Draw Uniform.
- Bayes calculation for memorylessness:

$$\mathbb{P}\{X > t + s \mid X > t\} = \frac{\mathbb{P}\{X > t + s \,\&\, X > t\}}{\mathbb{P}\{X > t\}} = \frac{\mathbb{P}\{X > t\}}{\mathbb{P}\{X > t\}}$$

- Random variables can be neither discrete nor continuous. Example: CDF with jump at 0, continuous thereafter.

3 Expectation

Definition 3. The expected value $\mathbb{E}[X]$ (mean) of a r.v. X is

- Discrete r.v.: $\mathbb{E}[X] = \sum_x x\mathbb{P}\{X = x\}$
- Continuous r.v.: $\mathbb{E}[X] = \int_x x f_X(x)dx$

If neither, other ways to define.

Note: Sometimes, these sums and integrals don't converge! $E[X]$ could be $\infty, -\infty$, or could be undefined.

3.1 Linearity of expectation

Scaling: If X is a r.v., and a is a real number, $\mathbb{E}[aX] = a\mathbb{E}[X]$

Adding: If X and Y are r.v.s, $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

3.2 Indicators and probabilities

$Bernoulli(p)$: 1 if heads, 0 if tails.

For event A, $\mathbb{1}\{A\}$: 1 if A happens, 0 if it doesn't happen.

$\mathbb{E}[\mathbb{1}\{A\}] = \mathbb{P}\{A\}$.

$\mathbb{E}[Bernoulli(p)] = p$

4 Variance

The deviation from average of r.v. is $Y := X - \mathbb{E}[X]$

Definition 4. The variance $\mathrm{Var}[X]$ of a r.v. X is defined to be $\mathrm{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[Y^2]$

Note: $\mathrm{Var}[X] \geq 0$, because it's squared.

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Remember: Always nonnegative! If it looks like it's negative, calculation mistake.

4.1 Scaling

$$\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$$
$$\mathrm{Var}[X+b] = \mathrm{Var}[X]$$

5 Summary of remainder

6 Joint Random Variables – 10 min

X is r.v., Y is also r.v.

- Joint CDF: $F_{X,Y}(x,y) = \mathbb{P}\{X \le x, Y \le y\}$
- If both discrete: Joint pmf: $\mathbb{P}\{X = x, Y = y\}$.
- If both continuous: Joint density: $f_{X,Y}(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}F_{X,Y}(x,y)$

"∂" is the partial derivative symbol, and it's pronounced "Partial".

Example: Urgent care waiting line at two different times – X = # of people waiting at noon, Y = # of people waiting at 1pm.

6.1 Marginals

Marginal just means to go back down to one r.v.

- CDF: Use limit: $\lim_{y\to\infty} F_{X,Y}(x,y) = F_X(x)$
- Density: Use integral: $\int_y f_{X,Y}(x,y)dy = f_X(x)$
- Pmf: Use sum: $\sum_y \mathbb{P}\{X = x, Y = y\} = \mathbb{P}\{X = x\}$

6.2 Independent Random Variables

Definition 5. X and Y are independent, $X \perp Y$, if

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)$$
$$\mathbb{P}\{X \le x, Y \le y\} = \mathbb{P}\{X \le x\}\mathbb{P}\{Y \le y\}$$

Useful fact:
For any functions g, h, if $X \perp Y$, then

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$$

For example: If $X \perp Y$,

$$\mathbb{E}[z^{X+Y}] = \mathbb{E}[z^X z^Y] = \mathbb{E}[z^X]\mathbb{E}[z^Y]$$

7 Covariance – 10 min

Definition 6. The covariance $Cov(X,Y)$ is:

$$\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

Alternatively: $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
Measures how correlated they are: Up at the same time, down at the same time.
Note: Can be negative, unlike variance.

7.1 Exercise: Covariance

This is important

Let's go back to our earlier example: X is# of people waiting in line at urgent care at noon, Y is at 1.

Say someone comes into urgent care, but the line's too long, so they want to come back later. We don't want the line to still be long when they get back. That's what Covariance measures! Let's calculate.

$X = 1, 3$, 1/2 each. $Y = X + 1, X, X - 1$, 1/3 each.

Let's use the formula $Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

What is $\mathbb{E}[XY]$?

$(1, 0), (1, 1), (1, 2), (3, 2), (3, 3), (3, 4)$

$\mathbb{E}[XY] = (0 + 1 + 2 + 6 + 9 + 12)/6 = 5$

$\mathbb{E}[X] = 2$

$\mathbb{E}[Y] = \mathbb{E}[X] + \mathbb{E}[\{-1, 0, 1, 1/3\ each] = \mathbb{E}[X] = 2$

$Cov(X, Y) = 5 - 4 = 1$

Is that low enough correlation? Covariance threshold to decide when patient should come back!