Singapore Prep Course

Probability, Chapter 2

Roger Lee

Expectation

Variance and covariance

Expectation

Definition: For discrete random variable X, the expectation of X is

$$\mathbb{E}(X) = \mathbb{E}X = \sum_x x\mathbb{P}(X = x)$$

(assuming absolute convergence for infinite sums).

- It is a mean, a weighted average of possible values, weighted by the probability of taking that value.
- If we draw independently many random variables $X_1, X_2, \dots$ with the same distribution as X, then

  $$\frac{X_1 + X_2 + \cdots + X_n}{n}$$

  should be close to $\mathbb{E}X$ with high probability.

  (This is made precise by the law of large numbers.)

Expectation examples

- If X is the result of a die roll then

$$\mathbb{E}X = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3.5$$

- Let A be any event. Let the random variable

$$\mathbf{1}_A := \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\mathbb{E}\mathbf{1}_A = 0 \times \mathbb{P}(A^c) + 1 \times \mathbb{P}(A) = \mathbb{P}(A)$$

Interview question

(JP Morgan)

If the probability of heads is p when tossing a coin, what is the average number of tosses needed to get a head?

Interview question

(JP Morgan)

A stock price which is currently 1.01. The stock can move only in steps of ± 0.01, and is a symmetric random walk. We will always put a limit order to buy 0.01 below the current price, thus 1.00 for now, if it goes down we get filled at 1.00, if it goes to 1.02 the new limit order is 1.01. What is the expectation of the price at which we finally buy the stock?

Expectation of continuous random variable

Definition: If X has density f then

$$\mathbb{E}X = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x$$

(assuming absolute integrability). Examples:

- If X is uniform on (a, b) then

$$\mathbb{E}X = \int_a^b x \frac{1}{b-a} \mathrm{d}x = \frac{1}{b-a} \frac{x^2}{2} \bigg|_a^b = \frac{a+b}{2}$$

- If X is Normal$(0, 1)$ then

$$\mathbb{E}X = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x = 0$$

- If X is Exponential(λ) then $\mathbb{E}X = \int_0^\infty x\lambda e^{-\lambda x} \mathrm{d}x = \frac{1}{\lambda}$

  Here λ is interpreted as a failure rate per unit time,

  and $1/\lambda$ is expected time until failure, or expected lifetime

Expectations of functions of random variables

If X is discrete then

$$\mathbb{E}g(X) = \sum_x g(x)\mathbb{P}(X = x)$$

If X is continuous with density f, then

$$\mathbb{E}g(X) = \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x$$

If (X, Y) are discrete then

$$\mathbb{E}g(X, Y) = \sum_x \sum_y g(x, y)\mathbb{P}(X = x, Y = y)$$

If (X, Y) are continuous with density f then

$$\mathbb{E}g(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y)\mathrm{d}x\mathrm{d}y$$

Interview question

(ING, online quant test)

Calculate $\mathbb{E}X^3$ where $X \sim \text{Normal}(\mu, \sigma^2)$.

Interview question

Two people, trying to meet, arrive at times independently and uniformly distributed between noon and 1pm. Find the expected length of time that the first waits for the second.

Interview question (DE Shaw)

A ten-floor building has two elevators. Suppose the positions of the elevators are independent and follow uniform distributions. On average it take 1 minute for you to wait, at the bottom floor. Today only one elevator is available. On average how long do you need to wait?

Linearity properties

For $a, b \in \mathbb{R}$ and random variables X, Y (regardless of independence):

$$\mathbb{E}(aX + b) = a\mathbb{E}X + b$$

and

$$\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y$$

because (in discrete case):

$$\sum_x \sum_y (x + y)\mathbb{P}(X = x, Y = y)$$

$$= \sum_x x \sum_y \mathbb{P}(X = x, Y = y) + \sum_y y \sum_x \mathbb{P}(X = x, Y = y)$$

$$= \sum_x x\mathbb{P}(X = x) + \sum_y y\mathbb{P}(Y = y) = \mathbb{E}X + \mathbb{E}Y$$

Interview question

(JP Morgan)

You have 18 people boarding 10 buses. What's the expected number of empty buses?

Interview question

(JP Morgan)

You have 10 blue balls and 8 red balls, lined up in a queue. What is the expected number of balls that have a neighbor of a different color?

Interview question

(Goldman Sachs)

I have a bag with 100 white balls. Randomly draw one. If it's white, mark it black. In any case, put it back. Repeat. What's the expected number of draws needed to mark all balls black?

Monotonicity and Multiplicativity

- If $X \geq Y$ with probability 1, then (regardless of indepedence)

$$\mathbb{E}X \geq \mathbb{E}Y$$

- If X, Y are independent, then

$$\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y)$$

because, in the discrete case,

$$\begin{aligned}\mathbb{E}(XY) &= \sum_x \sum_y xy\mathbb{P}(X = x, Y = y) = \sum_x \sum_y xy\mathbb{P}(X = x)\mathbb{P}(Y = y) \\ &= \sum_x x\mathbb{P}(X = x) \sum_y y\mathbb{P}(Y = y) = (\mathbb{E}X)(\mathbb{E}Y)\end{aligned}$$

Conditional Expectation

For a discrete X and an event A with $\mathbb{P}(A) > 0$, define the conditional expectation

$$\mathbb{E}(X|\ A) = \sum_x x\mathbb{P}(X = x|\ A).$$

Rule of total conditional expectation:

If we have a decomposition $\Omega = A_1 \cup A_2 \cup \cdots \cup A_n$ into disjoint events, then

$$\mathbb{E}X = \mathbb{E}(X|A_1)\mathbb{P}(A_1) + \mathbb{E}(X|A_2)\mathbb{P}(A_2) + \cdots + \mathbb{E}(X|A_n)\mathbb{P}(A_n)$$

Interview question

(Knight Capital)

Flip a coin until you get two heads consecutively. Let N be the number flips (including the two heads). Find $\mathbb{E}N$.

Interview question

(Goldman Sachs. Also Susquehanna.)

A bug is at the corner of a cube, what's the expectation of number of steps that the bug takes to go to the opposite corner? One step means the bug going from one vertex to its neighbor vertex.

Expectation

Variance and covariance

Variance

The variance of X is

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

It expresses how widely spread is the distribution, in the sense of having values far away from the mean.

- The two expressions are equivalent because, writing $\mu := \mathbb{E}X$, $\mathbb{E}[(X-\mathbb{E}X)^2] = \mathbb{E}(X^2-2\mu X+\mu^2) = \mathbb{E}X^2-2\mu\mathbb{E}X+\mu^2 = \mathbb{E}X^2-\mu^2$.
- $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$

Why squaring instead of taking absolute values $\mathbb{E}|X - \mathbb{E}X|$?

Because $\mathrm{Var}(X + Y) = \mathrm{Var}X + \mathrm{Var}Y$ for X and Y independent.

And squaring is differentiable, making $\min \mathrm{Var}(Y - (aX + b))$ easy.

Standard Deviation

Define the standard deviation of X by

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var} X}$$

Motivation: If X is in dollars, then $\mathrm{Var} X$ is in dollars-squared.

Take square root to recover dollars.

- So $\mathrm{Var}(\text{die roll}) = (2 \times 0.5^2 + 2 \times 1.5^2 + 2 \times 2.5^2)/6 = 35/12$ and $\mathrm{SD}(\text{die roll}) = \sqrt{35/12} \approx 1.7$.
- Standardizing a random variable: If Y has $0 < \mathrm{Var} Y < \infty$, then to *standardize* Y means to apply the transformation
$$Z := \frac{Y - \mathbb{E}Y}{\mathrm{SD}(Y)}$$
which translates and scales Y to have mean 0 and variance 1: $\mathbb{E}Z = \mathbb{E}(Y - \mathbb{E}Y)/\,\mathrm{SD}(Y) = 0$ and $\mathrm{Var} Z = (1/\,\mathrm{SD}(Y))^2 \mathrm{Var} Y = 1$.

Interview question

(JP Morgan)

In slide 2.5, find the variance of the number of tosses to get a head.

Covariance

Covariance of X and Y is

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X-\mathbb{E}X)(Y-\mathbb{E}Y)) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$$

Idea: if X and Y are high together and low together, Cov is positive, but if X is low when Y is high, and vice versa, then Cov is negative.

- The two expressions are equivalent because

$$\begin{aligned}\mathbb{E}[(X-\mu_X)(Y-\mu_Y)] &= \mathbb{E}(XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y) \\ &= \mathbb{E}(XY) - \mu_X\mu_Y\end{aligned}$$

  where $\mu_X := \mathbb{E}X$, $\mu_Y := \mathbb{E}Y$.

- $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

Correlation

If X and Y have nonzero finite variance, then define their correlation

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\mathrm{SD}(X)\,\mathrm{SD}(Y)}$$

So $\mathrm{Corr}(X, Y)$ has the same sign ($+$ or $-$ or 0) as $\mathrm{Cov}(X, Y)$.

- It can be shown that $-1 \leq \mathrm{Corr}(X, Y) \leq 1$.
- If $\mathrm{Corr}(X, Y) = 1$, then X and Y are "perfectly correlated". This is the case if $Y = aX + b$ where $a > 0$.
- If X and Y have zero covariance (or correlation), then X and Y are said to be uncorrelated.
- If X and Y are independent, then they are uncorrelated. (Interview question: if uncorrelated, are they necessarily independent?)

Examples

Consider the joint probability mass functions

Covariance

Bilinearity property

$$\mathrm{Cov}\Big(\sum_j a_j X_j, \sum_k b_k Y_k\Big) = \sum_j \sum_k a_j b_k \mathrm{Cov}(X_j, Y_k)$$

In particular:

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathrm{Cov}(X+Y, X+Y) \\
&= \mathrm{Cov}(X,X) + \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X) + \mathrm{Cov}(X,X) \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y).
\end{aligned}$$

So if X and Y are uncorrelated then $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Covariance

More generally,

$$\mathrm{Var}(X_1 + \cdots + X_n) = \sum_j \mathrm{Var}(X_j) + 2 \sum_{j,k} \mathrm{Cov}(X_j, X_k).$$

Compare:

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n)$$

regardless of whether $X_1, \ldots, X_n$ are correlated.

- The variance of a portfolio value is not determined by the individual variances, because it also depends on the correlations.
- The expectation of a portfolio value is determined by the individual expectations, regardless of the correlations.

Interview question (Goldman Sachs)

Let W_t be the position of some particle at time t.

Let W_{t_1} have mean 0 and variance t_1.

Let $W_{t_2} - W_{t_1}$ have mean 0 and variance $t_2 - t_1$.

Let W_{t_1} and $W_{t_2} - W_{t_1}$ be independent.

Find $\mathrm{Cov}(W_{t_1}, W_{t_2})$.