Expectations

March 13, 2019

Expectation

Discrete case

Suppose X is discrete RV. The expectation or the expected value of X is

$$
\begin{aligned}
E(X) &= \sum_{x \in Range(X)} xP(X = x) \\
&= \sum_{x \in Range(X)} x p_X(x)
\end{aligned}
$$

Meaning

- $E(X)$ is the weighted average value of X
- Think of X as a quantity defined by outcome of an experiment
- then if we repeat the experiment many times, the average value of X over all times will be about $E(X)$

Meaning

- $E(X)$ is the weighted average value of X
- Think of X as a quantity defined by outcome of an experiment
- then if we repeat the experiment many times, the average value of X over all times will be about $E(X)$

Meaning

- $E(X)$ is the weighted average value of X
- Think of X as a quantity defined by outcome of an experiment
- then if we repeat the experiment many times, the average value of X over all times will be about $E(X)$

Example

- Play a game, with probability $p(x_i)$ we win x_i dollars, $i = 1, \ldots, n$.
- play this game N times (N very large)
- Number of games win x_i is about $p(x_i)N$
- Average win per game

$$\frac{1}{N}\sum_{i=1}^{n} x_i p(x_i) N = E(X)$$

Example

I_A is the indicator RV for event A

$$I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

$$E(X) = 1P(A) + 0P(A^c) = P(A)$$

Example

I_A is the indicator RV for event A

$$I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

$$E(X) = 1P(A) + 0P(A^c) = P(A)$$

Continuous case

If X is continuous RV with pdf $f_X(x)$ then

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

Example

X has pdf

$$f(x) = \begin{cases} \frac{1}{2} & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \int_0^2 x\frac{1}{2}dx = 1$$

Properties

- If X is discrete RV then for real value function $g(x)$, $g(X)$ is RV and

$$E[g(X)] = \sum_x g(x)p_X(x).$$

- If X is continuous RV then

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

- Linear property

$$E(aX + b) = aE(X) + b$$

Example

X has pdf

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

then

$$E(X^3) = \int_0^1 x^3 dx = \frac{1}{4}$$

Example

X has pdf

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

then

$$E(X^3) = \int_0^1 x^3 dx = \frac{1}{4}$$

Expectation of joint RVs

X, Y had joint pmf $p_{X,Y}$ or joint pdf $f_{X,Y}$

- Discrete case

$$E(g(X,Y)) = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$$

- Continuous case

$$E(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dxdy$$

Expectation of joint RVs

X, Y had joint pmf $p_{X,Y}$ or joint pdf $f_{X,Y}$

- Discrete case

$$E(g(X,Y)) = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$$

- Continuous case

$$E(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$$

Expectation of Sum of RVs

Suppose X and Y are 2 random variables.
Then $X + Y$ is also a random variable and

$$E(X + Y) = E(X) + E(Y)$$

Proof

Find expectation for $g(X,Y) = X + Y$

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{X,Y}(x,y)dxdy$$

$$= \int_{-\infty}^{\infty} x(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy)dx$$

$$+ \int_{-\infty}^{\infty} y(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx)dy$$

$$= \int_{-\infty}^{\infty} xf_X(x)dx + \int_{-\infty}^{\infty} yf_Y(y)dy$$

$$= E(X) + E(Y)$$

In general, if $X_1, \ldots, X_n$ are RVs then

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

This is a very useful formula,
requires no assumption about X_i's

In general, if $X_1, \ldots, X_n$ are RVs then

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

This is a very useful formula, requires no assumption about X_i's

Example

A construction firm has recently sent in bids for 3 jobs worth (in profits) 10, 20, and 40 (thousand) dollars. If its probabilities of winning the jobs are respectively .2, .8, and .3, what is the firm's expected total profit?

Solution

Let X_i = profit from job i. The expected total profit is

$$E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3)$$
$$= [10(.2) + 0(.8)] + [20(.8) + 0(.2)]$$
$$+ [40(.3) + 0(.7)]$$
$$= 30$$

Example

A secretary has typed N letters along with their respective envelopes. The envelopes get mixed up when they fall on the floor. If the letters are placed in the mixed-up envelopes in a completely random manner (that is, each letter is equally likely to end up in any of the envelopes), what is the expected number of letters that are placed in the correct envelopes?

Solution

- X_i=indicator RV of {i-th letter in correct envelope}
- $E(X_i) = P(X_i = 1) = 1/N$
- Expected number of letters in correct envelope:

$$E(\sum_{i=1}^{N} X_i) = N\frac{1}{N} = 1$$

Solution

- X_i=indicator RV of {i-th letter in correct envelope}
- $E(X_i) = P(X_i = 1) = 1/N$
- Expected number of letters in correct envelope:

$$E(\sum_{i=1}^{N} X_i) = N\frac{1}{N} = 1$$

Prediction

- A RV X can have many values
- If we have to guess what value X would take, then $\mu = E(X)$ will be the best guess
- Suppose we guess $X = c$, the "average error" is
$$E(|X - c|^2)$$
- This error is minimized if $c = \mu$.

Prediction

- A RV X can have many values
- If we have to guess what value X would take, then $\mu = E(X)$ will be the best guess
- Suppose we guess $X = c$, the "average error" is
$$E(|X - c|^2)$$
- This error is minimized if $c = \mu$.

Prediction

- A RV X can have many values
- If we have to guess what value X would take, then $\mu = E(X)$ will be the best guess
- Suppose we guess $X = c$, the "average error" is
$$E(|X - c|^2)$$
- This error is minimized if $c = \mu$.

Proof

$$
\begin{aligned}
E[(X-c)^2] &= E[(X-\mu+\mu-c)^2] \\
&= E[(X-\mu)^2 + 2(\mu-c)(X-\mu) + (\mu-c)^2] \\
&= E[(X-\mu)^2] + 2(\mu-c)E[X-\mu] + (\mu-c)^2 \\
&= E[(X-\mu)^2] + (\mu-c^2) \quad \text{since} \quad E[X-\mu] = E[X]-\mu = 0 \\
&\geq E[(X-\mu)^2]
\end{aligned}
$$

Variance

- Given X and its pdf (or pmf)
- We want to have a ”feeling” of how X behaves
- The best guess of value of X is $\mu = E(X)$
- However μ alone is not enough

Variance

- Given X and its pdf (or pmf)
- We want to have a "feeling" of how X behaves
- The best guess of value of X is $\mu = E(X)$
- However μ alone is not enough

Variance

- Given X and its pdf (or pmf)
- We want to have a "feeling" of how X behaves
- The best guess of value of X is $\mu = E(X)$
- However μ alone is not enough

Variance

- Given X and its pdf (or pmf)
- We want to have a "feeling" of how X behaves
- The best guess of value of X is $\mu = E(X)$
- However μ alone is not enough

Variance

- We need another quantity to present how the values of X "spread" around μ
- Are the other values of X usually close to μ or can be far away?

Variance

- We need another quantity to present how the values of X ”spread” around μ
- Are the other values of X usually close to μ or can be far away?

Example

W, Y, Z: 3 RVs

$$W = 0 \quad \text{with probability } 1$$

$$Y = \begin{cases} -1 & \text{with probability } \frac{1}{2} \\ 1 & \text{with probability } \frac{1}{2} \end{cases}$$

$$Z = \begin{cases} -100 & \text{with probability } \frac{1}{2} \\ 100 & \text{with probability } \frac{1}{2} \end{cases}$$

Definition

Variance of X is the expectation of the square of difference between X and its mean μ.

$$\mathrm{Var}(X) = E(|X - \mu|^2)$$

or

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

$$
\begin{aligned}
\mathrm{Var}(X) &= E[(X-\mu)^2] \\
&= E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - E[2\mu X] + E[\mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&= E[X^2] - \mu^2
\end{aligned}
$$

Example

X = outcome of a fair dice

$$E(X) = \frac{1}{6}(1 + \cdots + 6) = \frac{21}{6}$$

$$E(X^2) = \frac{1}{6}(1^2 + \cdots + 6^2) = \frac{91}{6}$$

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{35}{12}$$

Practice

Find $\text{Var}(I_A)$ where I_A is the indicator RV of event A

$$I = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{if event } A \text{ does not occur} \end{cases}$$

Properties

- $\operatorname{Var}(X+b) = \operatorname{Var}(X)$
- $\operatorname{Var}(aX+b) = a^2 \operatorname{Var}(X)$
- $\operatorname{Var}(c) = 0$

Properties

- $\mathrm{Var}(X + b) = \mathrm{Var}(X)$
- $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$
- $\mathrm{Var}(c) = 0$

Properties

- $\mathrm{Var}(X+b) = \mathrm{Var}(X)$
- $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$
- $\mathrm{Var}(c) = 0$

Variance of sum

- Variance of sum of RVs are NOT the same as sum of the variances
- Ex:

$$\begin{aligned} \mathrm{Var}(X+X) &= \mathrm{Var}(2X) = 4\,\mathrm{Var}(X) \\ &\neq \mathrm{Var}(X) + \mathrm{Var}(X) \end{aligned}$$

- Need another quantity for multiple RVs

Variance of sum

- Variance of sum of RVs are NOT the same as sum of the variances
- Ex:

$$\begin{aligned}\text{Var}(X+X) &= \text{Var}(2X) = 4\,\text{Var}(X)\\ &\neq \text{Var}(X) + \text{Var}(X)\end{aligned}$$

- Need another quantity for multiple RVs

Variance of sum

- Variance of sum of RVs are NOT the same as sum of the variances
- Ex:

$$\begin{aligned}\text{Var}(X + X) &= \text{Var}(2X) = 4\ \text{Var}(X) \\ &\neq \text{Var}(X) + \text{Var}(X)\end{aligned}$$

- Need another quantity for multiple RVs

Practice

An insurance company writes a policy to the effect that an amount of money A must be paid if some event E occurs within a year. If the company estimates that E will occur within a year with probability p, what should it charge the customer so that its expected profit will be 10 percent of A?

Practice

Ten balls are randomly chosen from an urn containing 17 white and 23 black balls. Let X denote the number of white balls chosen. Compute $E[X]$

(a) by defining appropriate indicator variables $X_i, i = 1, \ldots, 10$ so that

$$X = \sum_{i=1}^{10} X_i$$

(b) by defining appropriate indicator variables $Y_i, = 1, \ldots, 17$ so that

$$X = \sum_{i=1}^{17} Y_i$$

Covariance

Definition

X and Y are two RVs. The **covariance** of X and Y is

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

Alternative formula

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

Definition

X and Y are two RVs. The **covariance** of X and Y is

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

Alternative formula

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

Meaning

- The covariance shows the relation between X and Y
- If $\mathrm{Cov}(X, Y) > 0$ then Y tends to increase when X increases
- If $\mathrm{Cov}(X < Y) < 0$ then Y tends to decrease when X increases

Meaning

- The covariance shows the relation between X and Y
- If $\operatorname{Cov}(X, Y) > 0$ then Y tends to increase when X increases
- If $\operatorname{Cov}(X < Y) < 0$ then Y tends to decrease when X increases

Meaning

- The covariance shows the relation between X and Y
- If $\operatorname{Cov}(X, Y) > 0$ then Y tends to increase when X increases
- If $\operatorname{Cov}(X < Y) < 0$ then Y tends to decrease when X increases

Properties

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$
- $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

Properties

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$
- $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

Properties

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$
- $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

Properties

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$
- $\mathrm{Cov}(X + Y, Z) =$
  $\mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

Variance of Sum

$X_1, \ldots, X_n$: RVs

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j)$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

Variance of Sum

$X_1, \ldots, X_n$: RVs

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i) + \sum_{i \neq j} \operatorname{Cov}(X_i, X_j)$$

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X, Y)$$

Independence

X and Y are independence if for all x, y

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

If so then

$$\mathrm{Cov}(X, Y) = 0$$

Proof

$$
\begin{aligned}
E[XY] &= \sum_j \sum_i x_i y_j P\{X = x_i, Y = y_j\} \\
&= \sum_j \sum_i x_i y_j P\{X = x_i\} P\{Y = y_j\} \quad \text{by independence} \\
&= \sum_y y_j P\{Y = y_j\} \sum_i x_i P\{X = x_i\} \\
&= E[Y]E[X]
\end{aligned}
$$

Corollary

If $X_i, \ldots, X_n$ are independent then

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Example

Compute the variance of number of heads from n independent tosses of a fair coin.

Solution

- X_i = indicator RV of { i-th toss is a head}
- $\sum X_i$ = total number of heads

$$\mathrm{Var}(\sum X_i) = \sum \mathrm{Var}(X_i)$$
$$= n\frac{1}{2}(1 - \frac{1}{2}) = \frac{n}{4}$$

Solution

- $X_i =$ indicator RV of { i-th toss is a head}
- $\sum X_i =$ total number of heads

$$\mathrm{Var}(\sum X_i) = \sum \mathrm{Var}(X_i)$$
$$= n\frac{1}{2}(1 - \frac{1}{2}) = \frac{n}{4}$$

Solution

- $X_i =$ indicator RV of { i-th toss is a head}
- $\sum X_i =$ total number of heads

$$\mathrm{Var}(\sum X_i) = \sum \mathrm{Var}(X_i)$$
$$= n\frac{1}{2}(1 - \frac{1}{2}) = \frac{n}{4}$$

Homework 5

Chapter 4: 12, 14, 21, 27, 38, 43, 45, 46