MATH 180A (Lecture A00)

mathweb.ucsd.edu/~ynemish/teaching/180a

Today: Expectation and variance of sums. Covariance and correlation

Next: ASV 9.1

Week 9:

- Homework 6 due Friday, March 10

Linearity of expectation

<u>Thm</u>. Let $X_1, \dots, X_n$ be random variables defined on the same probability space. Let $g_1, g_2, \dots, g_n$ be functions of one variable. Then

$$E(g_1(X_1) + g_2(X_2) + \cdots + g_n(X_n))$$

In particular, $E(X_1 + \cdots + X_n)$

Important: independence does not matter

Expectation of a sum = sum of expectations

ALWAYS!

Linearity of expectation

Example (Binomial distribution).

Let $X_1, \dots, X_n$ be independent random variables, $X_i \sim Ber(p)$

$$S_n = X_1 + \dots + X_n, \quad S_n \sim Bin(n,p).$$

$$E(S_n) =$$

Example Adam must pass both written test and road test for his driver's license. He passes written test with probability $\frac{4}{10}$, independently of other tests. For the road test, the probability of success is $\frac{7}{10}$. What is the total expected number of tests Adam must take before earning his license?

Denote $X = \#$ written tests before he passes

$Y = \#$ road tests before he passes

$$E(X+Y) - ?$$

Expectation of a product of independent random variables

Thm. Let $X_1, \dots, X_n$ be independent random variables.

Let $g_1, g_2, \dots, g_n$ be functions of one variable.

Then

$$E(g_1(X_1)\, g_2(X_2) \cdots g_n(X_n))$$

Corollary If $X_1, \dots, X_n$ are independent, then

Variance of a sum of independent random variables

Example

Binomial: $X_1, \dots, X_n$ independent identically distributed (iid)

$X_i \sim Ber(p)$, $Var(X_i) = p(1-p)$, $S_n = X_1 + \cdots + X_n$

$Var(S_n) =$

Sample mean: $X_1, \dots, X_n$ independent identically distributed (iid)

$E(X_i) = \mu$, $Var(X_i) = \sigma^2$

$E\left(\frac{X_1 + \cdots + X_n}{n}\right) =$, $Var\left(\frac{X_1 + \cdots + X_n}{n}\right) =$

Covariance

Suppose that we have a random variable X.

- E(X) - mean value, average of a large number of independent realizations
- Var(X) - variance, fluctuations of X, how far the realizations are spread around the mean

Covariance describes strength and type of dependence between two random variables.

Def. Let X and Y be random variables defined on the same probability space. The covariance of X and Y is given by

Computations:

Covariance

Example Let X, Y be discrete random variables with the joint PMF $P(X=k, Y=\ell)$ given by the table

$k \backslash \ell$	0	1	2	
-1	0.1	0	0.1	0.2
0	0.3	0.2	0	0.5
1	0	0.2	0.1	0.3
	0.4	0.4	0.2	

$E(X) =$

$E(Y) =$

$E(XY) =$

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

Some heuristics

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

- $(X-E(X))(Y-E(Y))$ is positive if

  $(X-E(X))$ and $(Y-E(Y))$ have the same sign

- $(X-E(X))(Y-E(Y))$ is negative if

  $(X-E(X))$ and $(Y-E(Y))$ have opposite signs

Thus,

- $\mathrm{Cov}(X,Y) > 0$ means that on average $X-E(X)$ and $Y-E(Y)$

  have

- $\mathrm{Cov}(X,Y) < 0$ means that on average $X-E(X)$ and $Y-E(Y)$

  have

- If $\mathrm{Cov}(X,Y) = 0$, we say that

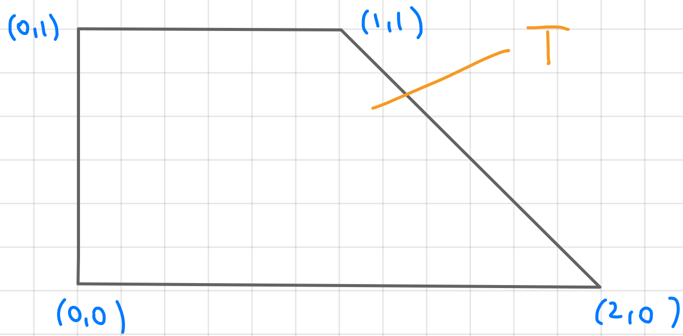
Example

Let (X, Y) be a uniformly distributed random point on the trapezoid with vertices $(0,0), (2,0), (1,1), (0,1)$

Is $\mathrm{Cov}(X,Y)$

(a) positive ?

(b) negative



Joint density: $f(x,y) = \frac{2}{3}$ for $(x,y) \in T$

$E(X) =$, $E(Y) =$

$E(XY) =$

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

Uncorrelated vs Independent

- X and Y are independent $\Rightarrow \operatorname{Cov}(X,Y)=0$

  $E(XY)$

- $\operatorname{Cov}(X,Y)=0 \not\Rightarrow X$ and Y are independent

Example of random variables X, Y that are <u>not</u> independent, but $\operatorname{Cov}(X,Y)=0$

Let $X \sim N(0,1)$, $Y = X^2$. Then

$$E(X)=0, \quad E(X^2)=1, \quad E(X^3)=0$$

$$\begin{aligned}\operatorname{Cov}(X,Y) &= E(XY) - E(X)E(Y) \\ &= E(X^3) - E(X)E(X^2) = 0\end{aligned}$$

Variance of a sum. Properties of covariance

Thm Let $X_1, \dots, X_n$ be random variables with finite variances.

Then $\quad \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) =$

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

Properties of covariance:

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

- $\mathrm{Cov}(aX+b, Y) =$

- $\mathrm{Cov}\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_i Y_i\right) =$

Correlation

Covariance is not particularly good for evaluating the strength of the dependence:

- suppose that $\mathrm{Cov}(X,Y)=1$, then $\mathrm{Cov}(10X,10Y)=100$, but the dependence between X and Y is the same as dependence between $10X$ and $10Y$.

Solution: normalize covariance $\longrightarrow$ correlation

Def. Let X, Y be random variables, $\mathrm{Var}(X)<\infty$, $\mathrm{Var}(Y)<\infty$

The correlation (coefficient) of X and Y is given by

Properties of correlation

Thm Let X, Y be random variables, $\mathrm{Var}(X) < \infty$, $\mathrm{Var}(Y) < \infty$

- $\mathrm{Corr}(aX+b, Y) =$
- $-1 \le \mathrm{Corr}(X,Y) \le 1$
- $\mathrm{Corr}(X,Y) = 1$ if and only if
- $\mathrm{Corr}(X,Y) = -1$ if and only if

Example Let X, Y be random variables satisfying

$$E(X) = 2, \quad E(Y) = 1, \quad E(X^2) = 5, \quad E(Y^2) = 10, \quad E(XY) = 1$$

(a) Compute $\mathrm{Corr}(X,Y)$

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

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

(b) Find $c \in \mathbb{R}$ such that X and $X + cY$ are uncorrelated.

Moment generating function of a sum of indep. RVs

Def. (Convolution of distributions - Section 7)

Let X and Y be random variables. Then the distribution of $X+Y$ is called the convolution of the distributions of X and Y.

If X and Y are continuous and f_X and f_Y are their PDFs then the PDF of $X+Y$ is given by

$$f_{X+Y}(s) = f_X * f_Y(s) = \int_{-\infty}^{+\infty} f_X(x) f_Y(s-x)\, dx = \int_{-\infty}^{+\infty} f_X(s-y) f_Y(y)\, dy$$

(similar formula for discrete random variables)

If X and Y are independent, it may be easier to compute the

Moment generating function of a sum of indep. RVs

Let X, Y be independent random variables.

Then the MGF of $X+Y$ is

1) $X \sim \text{Poisson}(\lambda)$, $Y \sim \text{Poisson}(\mu)$, independent.

Distribution of $X+Y$?

2) $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, independent.

Distribution of $X+Y$?