PROBABILITY THEORY
MM 3

MM 3: EXPECTATION AND VARIANCE

Topics:

Expectation and
Variance of a random variable.
Covariance and Variance of
sums of random variables.

Post-mm2 questions

Q1. If we know marginal pdfs of random variables X and Y, are we able to find their joint pdf?

Q2. Let X be a continuous r.v. Calculate P(X=100). $= 0$

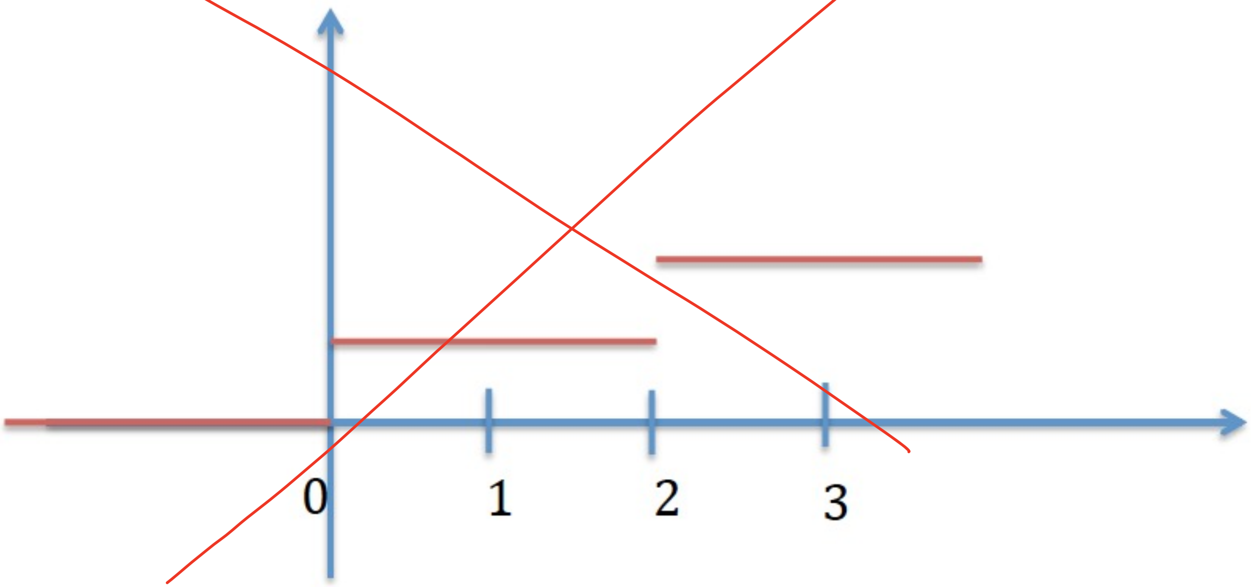
Q1 marginal pdf $\xrightarrow[?]{}$ joint pdf

X , Y

$\nrightarrow$

$\leftarrow$

Q3. Cmf of r.v. X is given on a figure below. Find P(X=0), P(X=1), P(X=2), P(X=3).



Introduction

- In order to completely describe the behavior of a r.v. an entire function – the cdf or pdf – must be given. In some situations we are interested just in a few parameters that summarize the information provided by these functions.
- For example, when a large collection of data is assembled, we are typically interested not in the individual numbers, but rather in a certain quantities such as the average.

A r.v. X

pmf (discrete)

pdf (cont.)

cdf

Calculate

$P(X = 100)$

$P(X \in (0,1))$

$P(X > 3)$

$X \in (3, \infty)$

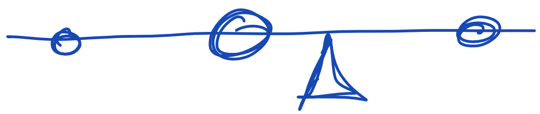
"Key performance indicators"

- Expectation
- Variance

What should we learn today?

- How to calculate expectation and variance of a r.v.?
- How to calculate the expectation and variance of a sum of random variables?
- What is a moment generating function?

Expectation



- Expectation = expected value = mean = mean value
- Definition. The expectation of X is defined by (if X is a discrete r.v.)

$$E[X] = \sum_i x_i P(X = x_i) = \sum_i x_i p(x_i)$$

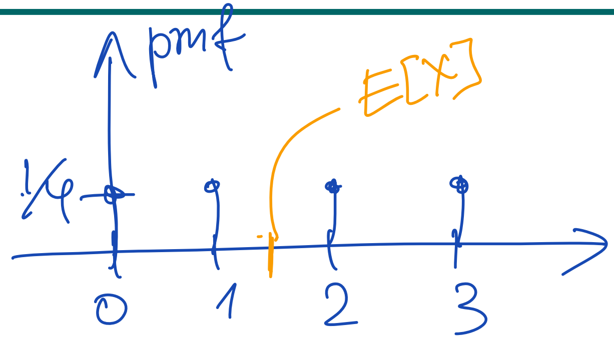
$X \in \{x_i\}$ weighted sum

- The expectation of X is a weighted average of the possible values that X can take on
- If X is a continuous r.v.

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

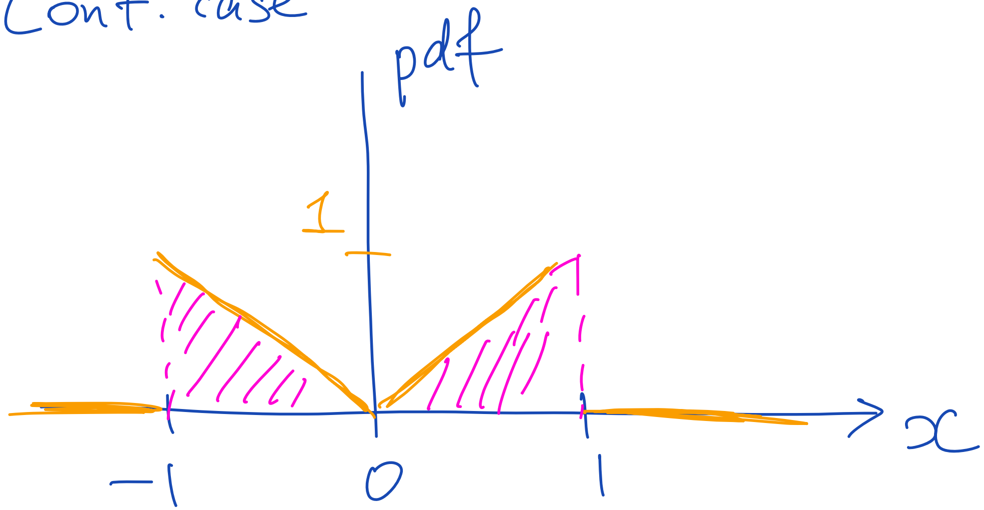
Ex

$X \in \{0, 1, 2, 3\}$



$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} = 1{,}5$$

Cont. case



$$f(x) = \begin{cases} x, & x \in (0,1) \\ -x, & x \in (-1,0) \\ 0, & \text{otherwise} \end{cases}$$

$$\boxed{\int_{-\infty}^{\infty} f(x)\,dx = 1}$$

$$E[X] = \int_{-1}^{1} f(x) \cdot x\,dx = \int_{-1}^{0} (-x) \cdot x\,dx + \int_{0}^{1} x \cdot x\,dx = 0$$

Expectation

- If instead of density function we consider distribution function, we can use the following expressions in case when a discrete distribution taking values 0, 1, 2, … and continuous distribution is taking nonnegative values :

$$E[X] = \sum_{k=0}^{\infty} P(X > k)$$

$$E[X] = \int_0^{\infty} (1 - F(x))dx$$

- The expected value is defined if the above sum or integral converges absolutely:

$$\sum_i |x_i| p(x_i) < \infty$$

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

- Example: the expectation does not exist for

$$F(x) = 1 - \frac{1}{x}, \quad x > 1$$

- Physical analogy: centre of gravity; E[X] has the same units of measurements as X

$$\mathrm{E}[X] = \sum_{n=0}^{\infty} \mathrm{P}(X > n).$$

This formula is valid only for a non-negative integer-valued random variable

$$\sum_{n=0}^{\infty} \mathrm{P}(X > n) = \sum_{n=0}^{\infty} \sum_{j=n+1}^{\infty} \mathrm{P}(X = j).$$

$$M = \begin{bmatrix} \mathrm{P}(X=1) & \mathrm{P}(X=2) & \mathrm{P}(X=3) & \cdots & \mathrm{P}(X=n) & \cdots \\ & \mathrm{P}(X=2) & \mathrm{P}(X=3) & \cdots & \mathrm{P}(X=n) & \cdots \\ & & \mathrm{P}(X=3) & \cdots & \mathrm{P}(X=n) & \cdots \\ & & & \ddots & \vdots & \\ & & & & \mathrm{P}(X=n) & \cdots \\ & & & & & \ddots \end{bmatrix}.$$

$$\mathbf{E}[X] = \int_{\overline{\mathbb{R}}} x \, dF(x),$$

$$\mathrm{E}[X] = \int_0^{\infty} (1 - F(x)) \, dx - \int_{-\infty}^{0} F(x) \, dx,$$

Expectation of a function of a random variable

X; new r.v. $Y = g(X)$, e.g. $Y = X^2$; $Y=2X+1$, $Y=e^X$

- Suppose that we are given a r.v. X and its probability distribution. Suppose we are interested in finding not the expected value of X, but the expected value of some function of X, g(X).
- g(X) is itself a r.v. with some probability distribution.
- How to obtain E[g(X)]?
  - obtain distribution for g(X) and use definition

$$\boxed{E[Y] = E[g(X)] = \sum_i g(x_i)p(x_i)}$$

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

X, pmf X $\longrightarrow$ $Y = g(X)$; calculate pmf Y + use definition of $E[Y]$

$$X = \begin{cases} 2, & p = 1/4 \\ 1, & \text{—"—} \\ -1, & \\ -2, & \end{cases}$$

$X \in \{-2, -1, 1, 2\}$

$$Y = X^2 = \begin{cases} 4, & p = 1/2 \\ 1, & p = 1/2 \end{cases}$$

"compression" step

$$\boxed{Y = X^2 = \begin{cases} 4, & p = 1/4 \\ 1, & p = 1/4 \\ 1, & p = 1/4 \\ 4, & p = 1/4 \end{cases}}$$

$$E[Y] = 1 \cdot \frac{1}{2} + 4 \cdot \frac{1}{2} = 2{,}5$$

"Direct" calculation:

$$E[Y] = 4 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4}$$

LOTOS

$$\boxed{E[Y] = E[g(X)] = \sum_{\text{all } i} g(x_i) \cdot p(x_i)}$$

$Y = 2X - 1$ $\quad E[Y] = (2 \cdot 2 - 1) \cdot \frac{1}{4} + (1 \cdot 2 - 1) \cdot \frac{1}{4} + (-1 \cdot 2 - 1) \cdot \frac{1}{4} + (-2 \cdot 2 - 1) \cdot \frac{1}{4} = \ldots$

$= 2 \cdot 0 - 1 = -1$ $E[2X-1] = 2E[X] - 1 = 2 \cdot 0 - 1 = -1$

Expectation of a function of a random variable

$Y = g(X)$

r. v.	→ values	→ respective probabilities
X	x_i	$p(x_i)$
Y	$g(x_i)$	$p(x_i)$

$$E[X] = \sum_i x_i \, p(x_i)$$

$$E[Y] = \sum_i g(x_i) \, p(x_i)$$

Properties of the expected value

Proposition:

$$\boxed{E[aX + b] = aE[X] + b}$$

$$E[b] = b$$

Intuitive that the expestation of a constant is the constant itself

$$E[aX] = aE[X]$$

$$E[X + b] = E[X] + b$$

We can shift the mean of a r.v.
By adding a constant to it

Proof (for cont. case) $\quad Y = g(X) = aX + b$

$$E[aX+b] = \int_{-\infty}^{\infty} (ax+b) f(x)\,dx =$$

$$= \int_{-\infty}^{\infty} \left(axf(x) + bf(x) \right) dx = \int_{-\infty}^{\infty} axf(x)\,dx + \int_{-\infty}^{\infty} bf(x)\,dx =$$

$$= a \underbrace{\boxed{\int_{-\infty}^{\infty} xf(x)\,dx}}_{E[X]} + b \underbrace{\boxed{\int_{-\infty}^{\infty} f(x)\,dx}}_{=1} = aE[X] + b$$

Proof of linear property of expectation

Discrete case:

$$E[aX+b] = \sum_i (ax_i + b)\,p(x_i) =$$

$$= a\sum_i x_i\,p(x_i) + b\sum_i p(x_i) = a\,E[X] + b$$

Continuous case:

$$E[aX+b] = \int_{-\infty}^{+\infty} (ax+b)\,f(x)\,dx =$$

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

Expected value of function of random variables

- Many important results in probability theory concerns sums of random variables.
- We first consider, what it means to add two r.v.

- In general case, we consider a function of two random variables Z=g(X,Y). The problem of finding the mean of Z=g(X,Y) is similar to the problem of finding the mean of a function of a single r.v.:

$$\boxed{E[Z] = E[g(X,Y)] = \sum_i \sum_j g(x_i, y_j) p(x_i, y_j)}$$

$$\boxed{E[Z] = E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) dx dy}$$

X, Y

$p(x_i, y_j)$ joint pmf

$f(x, y)$ joint pdf

$$Z = g(X, Y)$$

$$\boxed{Z = X + Y}$$

$$Z = X \cdot Y$$

$$Z = 2X + X^Y + \sin\left(\frac{X}{Y}\right)$$

$$\left.\begin{array}{l} X = x_i \\ Y = y_j \end{array}\right\} \rightarrow Z = g(x_i, y_j)$$

$$\boxed{\begin{array}{l} E[Z] = \sum_{\text{all } i}\sum_{j} g(x_i, y_j) \cdot p(x_i, y_j) \\ E[Z] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\,dx\,dy \end{array}}$$

calculation of E without first finding f_Z (p_Z)

Expectation of the sum of r.v.

- For example, if g(X,Y)=X+Y

$$x f(x,y) + y f(x,y)$$

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

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dxdy =$$

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

$$E[X] + E[Y]$$

$$\int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x,y)dy\right] dx$$

$$= f_X(x)$$

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

- It is easy to prove by mathematical induction, that the expected value of the sum of any finite number of random variables is the sum of the expected values of the individual random variables

- Note, that the r.v. do not have to be independent; formula is valid for any kind of r.v.

$$E[X+Y] = E[X] + E[Y]$$

ALWAYS !!!

HERRA

the most fantastic fact

Expectation of the product of r.vs.

$E[X \cdot Y] = E[X] \cdot E[Y]$; ONLY X, Y are independent

$Z = X \cdot Y$ $\qquad Z = g_1(X) \cdot g_2(Y)$

more generally, we consider $Z = g(X,Y) = g_1(X) \cdot g_2(Y)$

$$E[Z] = E[g_1(X) \cdot g_2(Y)] =$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_1(x)\, g_2(y) \underbrace{f(x,y)}_{= f_X(x) \cdot f_Y(y)} dx\, dy$$

if X and Y are independent

$$= \int_{-\infty}^{+\infty} g_1(x) f_X(x)\, dx \cdot \int_{-\infty}^{+\infty} g_2(y) f_Y(y)\, dy = E[g_1(X)]\, E[g_2(Y)]$$

$$\boxed{E[X \cdot Y] = E[X] \cdot E[Y],}$$

ONLY IF X and Y are independent

Prediction of a value of a r.v.

- The best predictor of a r.v., in terms of minimizing its mean square error, is its expectation.

for any constant c

$$E[(X-c)^2] = E[(X-\mu+\mu-c)^2] =$$

$$= E[(X-\mu)^2] + 2(\mu-c)E[X-\mu] + (\mu-c)^2 \geq$$

$$\geq E[(X-\mu)^2]$$

since $E[X-\mu] = E[X] - \mu = 0$

Variance

- Motivation: The variations are given by r.v. D=X-E[X]. D can take on positive and negative values. We are only interested in magnitude → we work with |D| or D^2.
- Definition: variance of r.v. X is defined as

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

X – meters
$E[X]$ – meters
$Var(X)$ – m^2
$std(X)$ – meters

- This expression can be simplified:

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

- Definition: Standard deviation of r.v. X is $\boxed{Std(X) = \sqrt{(Var(X))}}$
- Example: Variance of an Indicator of an event

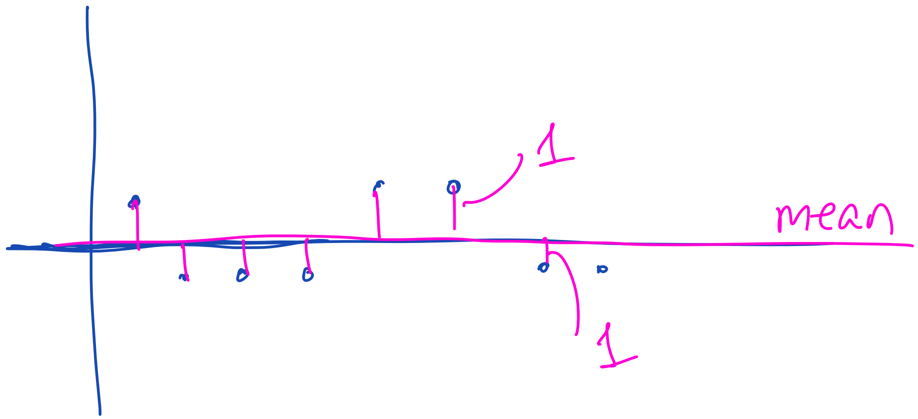
$$X = \begin{cases} 1 & , \ 1/2 \\ -1 & , \ 1/2 \end{cases}$$

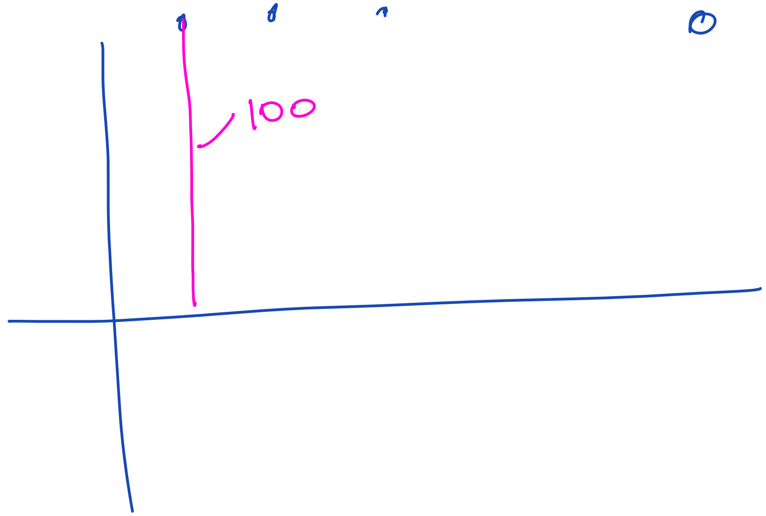
$$Y = \begin{cases} 100 & , \ 1/2 \\ -100 & , \ 1/2 \end{cases}$$

$$E[X] = 0$$

$$E[Y] = 0$$

realization



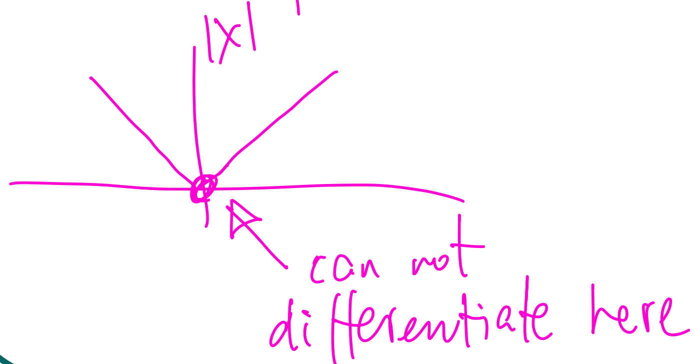


$$|X - E[X]| \to (X - E[X])^2$$ a r.v.

$$\downarrow$$

number

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



Another formula for variance

$$E[X^2 - 2XE[X] + E[X]^2] =$$

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

$$= E[X^2 - 2E[X]X + E[X]^2] =$$

$$= E[X^2] + E[-2E[X]X] + E[E[X]^2]$$

$$= E[X^2] - 2E[X]E[X] + E[X]^2 =$$

$$= E[X^2] - 2E[X]^2 + E[X]^2 =$$

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

$$E[-2E[X]X] = -2E[X]E[X] = \underline{-2E[X]^2}$$

$$E[E[X]^2] = E[X]^2$$

Variance

- Some useful identities concerning variances:

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

$$Var(b) = 0$$

$$Var(X + b) = Var(X)$$

Derivation of formula for *Var(aX+b)*

$$Var(aX+b) = E[(aX+b-E[aX+b])^2] =$$

$$= E[(aX+b-aE[X]-b)^2] = E[(aX-aE[X])^2] =$$

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

Variance of a sum of random variables

- Question:
- We have seen that the expectation of a sum of r.v. is equal to the sum of their expectation.
- Is the corresponding statement for variances true?

$$Var(X+Y) = Var(X) + Var(Y) +$$

+ something more

- No, it's not true. Example showing this:

$$\mathrm{Var}(X+X) = \mathrm{Var}(2X) = 4\,\mathrm{Var}(X) \neq \mathrm{Var}(X) + \mathrm{Var}(X)$$

- If r.vs. are independent, then it is true!

Covariance

- Definition. Covariance of 2 r.v. X and Y is defined by

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

- This expression can be simplified:

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

$$Cov(X,X) = E[(X - E(X)) \cdot (X - E(X))] = Var(X)$$

Simplification of the expression for covariance

$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])] =$$

$$= E[XY - XE[Y] - YE[X] + E[X]E[Y]] =$$

$$= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] =$$

$$= E[XY] - E[X]E[Y]$$

$$E[-XE[Y]] = -E[Y] \cdot E[X]$$

Properties of covariance

1. $Cov(X,Y) = Cov(Y,X)$
2. $Cov(X,X) = Var(X)$
3. $Cov(aX,Y) = aCov(X,Y)$
4. $Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$
5. If X and Y are independent, $Cov(X,Y) = 0$
6. $\boxed{Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)}$

   (thus, $Var(X+Y) = Var(X) + Var(Y)$ only in case of independent r.v.)

Property no 4 - proof

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

$$= \underline{E[XZ]} + \underline{E[YZ]} - \underline{E[X]E[Z]} - \underline{E[Y]E[Z]} =$$

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

- Generalizing, using induction principle we can prove for any *n* and any *m*:

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

Property no 5 - proof

$$Cov(X,Y) = E[\underbrace{(X-E[X])}_{g_1(X)}\underbrace{(Y-E[Y])}_{g_2(Y)}] =$$

$$= E[g_1(X)]\cdot E[g_2(Y)] = \underbrace{E[X-E[X]]}_{=0}\cdot\underbrace{E[Y-E[Y]]}_{=0} = 0$$

$$E[X-E[X]] = E[X]-E[X] = 0$$

$$E[X-E[X]] = E[X]-E[E[X]] = E[X]-E[X] = 0$$

- If X and Y are independent, then covariance is zero (X and Y are uncorrelated).
- It is possible that X and Y are uncorrelated, but not independent.

X and Y are independent $\Longrightarrow$ $\mathrm{Cov}(X,Y)=0$

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

When $\mathrm{Cov}(X,Y)=0$, $\Rightarrow$ X and Y are uncorrelated

What shows covariance?

- Covariance is a measure of the joint variability of 2 r.vs.

- If two variables tend to vary together 89when one os above its expectation, then the other is also above its expectation), then covariance is positive. And negative otherwise.

What shows covariance? (cntd)

if A and B are independent, $P(AB) = P(A)P(B)$ and
$Cov(X,Y) = 0$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = P(AB) - P(A)P(B)$$

$$Cov(X,Y) > 0 \iff P(AB) > P(A)P(B) \iff \frac{P(AB)}{P(A)} > P(B)$$

$$\iff P(B|A) > P(B)$$

Variance of a sum of r.vs.

Correlation coefficient

- Instead of covariance, we can work with dimensionless quality:
- Correlation coefficient of X and Y

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{(Var(X))}\sqrt{(Var(Y))}}$$

- Values of correlation coefficient lie between –1 and 1.

$-1 < Corr < 1$

if $Corr = 1, \Rightarrow X = Y$

if $Corr = -1 \Rightarrow X = -Y$

The nth moment of r.v. X

- The mean value is sometimes called the first moment of X.
- Definition. The nth moment of X is defined by

$$\boxed{E[X^n] = \sum_i x_i^n p(x_i)}$$

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

Moment generating function

- Definition. The moment generating function of a r.v. X is defined for all values t by

$$\boxed{\varphi(t) = E[e^{tX}] = \sum_i e^{tx_i} p(x_i)}$$

$$\boxed{\varphi(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx}$$

- This function is called moment generating function, because all moments of X can be obtained by successively differentiating it.
- There is a one-to-one correspondence between the moment generating function and the distribution function of a r.v.: the mgf uniquely determines the distribution and vice versa.

How to obtain E[X] and Var(X) from moment generating function?

$$\varphi'(t) = \frac{d}{dt} E[e^{tX}] = E\left[\frac{d}{dt}(e^{tX})\right] =$$

$$= E[X e^{tX}]$$

$$\varphi'(0) = E[X]$$

$$\varphi''(t) = \frac{d}{dt} E[X e^{tX}] = E\left[\frac{d}{dt}(X e^{tX})\right] =$$

$$= E[X^2 e^{tX}]$$

$$\varphi''(0) = E[X^2]$$

$$Var(X) = \varphi''(0) - \varphi'(0)^2$$

Example: packet retransmissions

$$P\{M = k\} = (1-p)^{k-1}p$$

$$E[M] = \sum_{k=1}^{+\infty} k\,p\,(1-p)^{k-1}$$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

$$\sum_{k=0}^{+\infty} k\,x^{k-1} = \frac{1}{(1-x)^2}$$

$$x = 1-p$$

$$E[M] = p \cdot \frac{1}{(1-(1-p))^2} = \frac{p}{p^2} = \frac{1}{p}$$

$$E[M] = \frac{1}{p} \qquad Var(M) = \frac{1-p}{p^2}$$

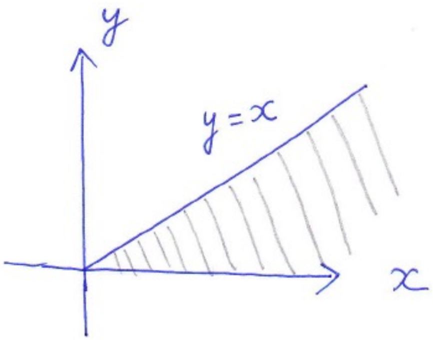
M	Probability
1	p
2	$(1-p)p$
3	$(1-p)^2p$
⋮	
k	$(1-p)^{k-1}p$
⋮	

M^2	Pr
1	p
2^2	$(1-p)p$
3^2	$(1-p)^2p$
⋮	
k^2	$(1-p)^{k-1}p$

Example: covariance

joint pdf for X and Y:

$$f_{X,Y}(x,y) = \begin{cases} c e^{-x} e^{-y}, & 0 \le y < x < \infty \\ 0, & \text{otherwise} \end{cases}$$



The constant c is found from the normalization condition:

$$1 = \int_0^{+\infty} \int_0^{x} c e^{-x} e^{-y} \, dy \, dx = \frac{c}{2} \Rightarrow c = 2$$

Example: covariance (ctnd)

The marginal pdfs:

$$f_X(x) = \int_0^x 2e^{-x}e^{-y}\,dy = 2e^{-x}(1-e^{-x}), \quad x \in [0,\infty)$$

$$f_Y(y) = \int_0^\infty 2e^{-x}e^{-y}\,dx = 2e^{-2y}, \quad y \in [0,\infty)$$

$$E[X] = \frac{3}{2}, \quad Var(X) = \frac{5}{4}$$

$$E[Y] = \frac{1}{2}, \quad Var(Y) = \frac{1}{4}$$

Example: covariance (ctnd)

$$E[XY] = \int_0^\infty \int_0^x xy \cdot 2e^{-x}e^{-y}\,dy\,dx =$$

$$= \int_0^\infty 2xe^{-x}(1-e^{-x}-xe^{-x})\,dx = 1$$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = 1 - \frac{3}{4} = \frac{1}{4}$$

$$Corr(X,Y) = \frac{1/4}{\sqrt{5/4}\sqrt{1/4}} = \frac{1}{\sqrt{5}}$$