EX 25: Let $X \sim U(0,2)$ and $Y = X^2+1$

1. Find the distribution function of the random variable Y.
2. Find $cov(X,Y)$
3. Are the random variables X and Y independent? Why?

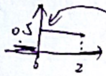
1. The distribution function of X is

$P(X\le ?) = F(?)$

$P(X \le \delta) = F(\delta)$

$$F_X(x) = 0 \quad x<0$$
$$= \frac{x}{2} \quad 0<x\le 2$$
$$= 1 \quad x>2$$

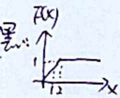
2. $f(x) = \begin{cases} 0 & x\le 0,\ x>2 \\ \frac{1}{2} & 0<x\le 2 \end{cases}$



the distribution function of Y is then

1. $F_Y(y) = P(Y\le y) = P(X^2+1\le y) = P(X \le \sqrt{y-1}) = F_X(\sqrt{y-1})$

i.e. $F_Y(y) = 0 \quad y<1$
$$= \frac{\sqrt{y-1}}{2} \quad 1\le y\le 5$$
$$= 1 \quad y>5$$



$F(x) = \begin{cases} 0 & x<0 \\ \frac{x}{2} & 0\le x\le 1 \\ 1 & x\ge 1 \end{cases}$

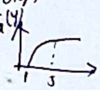
$\parallel$

$P(X\le x)$

$G(y) = P(Y\le y) = P(X^2+1\le y) = P(X\le\sqrt{y-1}) = F(\sqrt{y-1}) = \frac{\sqrt{y-1}}{2}$ for $1\le y\le 5$

$0 \quad y<1$

$1 \quad y>5$



2. $cov(X,Y) = EXY - EX\cdot EY = EX(X^2+1) - EX\,E(X^2+1)$
$= EX^3 + EX - EX\,EX^2 - EX = EX^3 - EX\,EX^2$

$EX = \int_0^2 x\left(\frac{1}{2}\right)dx = 1$

$EX^2 = \int_0^2 x^2\left(\frac{1}{2}\right)dx = \frac{4}{3}$

$EX^3 = \int_0^2 x^3\left(\frac{1}{2}\right)dx = 2$

Thus $cov(X,Y) = 2 - 1\cdot\frac{4}{3} = \frac{2}{3}$

3. Not independent, $\Leftarrow cov(X,Y)\ne 0$.

EX 26: Joint probability of the random variables X and Y is given by the following table:

	$X=0$	$X=1$	$X=2$
$Y=0$	$1/4$	$1/8$	0
$Y=1$	$1/4$	$1/4$	$1/8$

1. Find their marginal distributions
2. Find their covariance matrix and correlation matrix
3. Are the random variables X and Y independent? why?

1. The marginal distribution of the random variable X is.

$P(X=0) = P(X=0, Y=0) + P(X=0, Y=1) = \frac{1}{4}+\frac{1}{4} = \frac{1}{2}$

$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) = \frac{1}{8}+\frac{1}{4} = \frac{3}{8}$

$P(X=2) = P(X=2, Y=0) + P(X=2, Y=1) = \frac{1}{8}+0 = \frac{1}{8}$

The marginal ... of Y is.

$P(Y=0) = \frac{1}{4}+\frac{1}{8}+0 = \frac{3}{8}$

$P(Y=1) = \frac{1}{4}+\frac{1}{4}+\frac{1}{8} = \frac{5}{8}$

3. Not independent. $\Rightarrow cov(X,Y)\ne 0$.

Another reason:

$P(X=0, Y=0) = \frac{1}{4} \ne P(X=0).P(Y=0) = \frac{1}{2}\cdot\frac{3}{8}$

2. $cov(X,Y) = EXY - EX\cdot EY$

$EX = 0\cdot\frac{1}{2} + 1\cdot\frac{3}{8} + 2\cdot\frac{1}{8} = \frac{5}{8}$

$EY = 0\cdot\frac{3}{8} + 1\cdot\frac{5}{8} = \frac{5}{8}$

$EXY = 0\cdot0\cdot\frac{1}{4} + 0\cdot1\cdot\frac{1}{4} + 1\cdot0\cdot\frac{1}{8} + 1\cdot1\cdot\frac{1}{4} + 2\cdot0\cdot0 + 2\cdot1\cdot\frac{1}{8} = \frac{1}{2}$

$cov(X,Y) = EXY - EX\cdot EY = \frac{1}{2} - \frac{5}{8}\cdot\frac{5}{8} = \frac{7}{64}$

$EX^2 = 0^2\cdot\frac{1}{2} + 1^2\cdot\frac{3}{8} + 2^2\cdot\frac{1}{8} = \frac{7}{8}$

$EY^2 = 0^2\cdot\frac{3}{8} + 1^2\cdot\frac{5}{8} = \frac{5}{8}$

$var X = EX^2 - (EX)^2 = \frac{7}{8} - \left(\frac{5}{8}\right)^2 = \frac{31}{64}$

$var Y = EY^2 - (EY)^2 = \frac{5}{8} - \left(\frac{5}{8}\right)^2 = \frac{15}{64}$

The covariance matrix is then

$$cov(X,Y) = \begin{pmatrix} \frac{31}{64} & \frac{7}{64} \\ \frac{7}{64} & \frac{15}{64} \end{pmatrix} = \begin{pmatrix} var X & cov(X,Y) \\ cov(Y,X) & var Y \end{pmatrix}$$

$$corr(X,Y) = \frac{cov(X,Y)}{\sqrt{var X}\cdot\sqrt{var Y}} = \frac{7/64}{\sqrt{31/64}\sqrt{15/64}} = \frac{7}{\sqrt{465}}$$

, so correlation matrix is $\rho(X,Y) = \begin{pmatrix} 1 & \frac{7}{\sqrt{465}} \\ \frac{7}{\sqrt{465}} & 1 \end{pmatrix}$

Ex 7: Joint density of the random
$f_{(X,Y)}(x,y) = \begin{cases} \frac{1}{2}e^{-x-\frac{y}{2}}, & x>0, y>0 \\ 0 & \text{otherwise} \end{cases}$

1. Find their marginal distributions
2. Are the random variables X and Y independent? Why?
3. Find their covariance matrix and correlation matrix.

1. $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_0^{\infty} \frac{1}{2}e^{-x-\frac{y}{2}}\,dy = \frac{1}{2}e^{-x}\cdot[-2e^{-\frac{y}{2}}]_0^{\infty} = e^{-x}$ for $x>0$.
$= 0$ otherwise

$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_0^{\infty} \frac{1}{2}e^{-x-\frac{y}{2}}\,dx = \frac{1}{2}e^{-\frac{y}{2}}\cdot[-e^{-x}]_0^{\infty} = \frac{1}{2}e^{-\frac{y}{2}}$ for $y>0$,
$= 0$ otherwise.

2. they are indep iff $f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y)$, $\forall x,y$. from 1 we know that satisfied

3. Since $X \sim Exp(1)$ and $Y \sim Exp(\frac{1}{2}) \Rightarrow Var\,X = 1$ and $Var\,Y = 4$.
From the independence of X and Y, it follows that $cov(X,Y)=0 \Rightarrow corr(X,Y)=0 \Rightarrow$
$x \sim Exp(\lambda)$
$Var\,x = \frac{1}{\lambda^2}$

$Cov(X,Y) = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ and $\rho(X,Y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

EX.

Y \ X	-1	0	1	
1	1/9	1/9	1/9	1/3
2	1/3	0	1/3	2/3
	4/9	1/9	4/9	

a). Calculate COV(X,Y)
b). Are X, Y independent? why?
c). Find the joint distribution (i.e. the table in this case) of U and V.

a). $EX = -\frac{4}{9} + 0 + \frac{4}{9} = 0$
$EY = 1\cdot\frac{1}{3} + 2\cdot\frac{2}{3} = \frac{5}{3}$
$EXY = 1\cdot(-1)\cdot\frac{1}{9} + 1\cdot 0\cdot\frac{1}{9} + 1\cdot 1\cdot\frac{1}{9} + (-1)\cdot 2\cdot\frac{1}{3} + 1\cdot 2\cdot\frac{1}{3} = 0$
$cov(X,Y) = E(XY) - EXEY = 0$

b). No.
$P(X=-1, Y=1) \neq P(X=-1)\cdot P(Y=1)$
$\frac{1}{9} \neq \frac{4}{9}\cdot\frac{1}{3} \neq \frac{1}{9}$.

c). Find the joint distribution (i.e the table in this case) of U and V, where
U has the same marginal distribution as X.
V ——— 〃 ——— Y.

Y \ X	-1	0	1	
1	4/27	1/27	4/27	1/3
2	8/27	2/27	8/27	2/3
	4/9	1/9	4/9	

Ex 7: (1). $f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dy = \int_{-\infty}^{0} 0\,dy + \int_0^{\infty} \frac{1}{2}e^{-x-\frac{y}{2}}\,dy = \begin{cases} e^{-x} & x>0 \\ 0 & x<0 \end{cases}$

$f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dx = \int_{-\infty}^{0} 0\,dy + \int_0^{\infty} \frac{1}{2}e^{-x-\frac{y}{2}}\,dx = \begin{cases} \frac{1}{2}e^{-\frac{y}{2}} & y>0 \\ 0 & y<0 \end{cases}$

$\Rightarrow X \sim Exp(1)$
$Y \sim Exp(\frac{1}{2})$

$Z \sim Exp(\lambda)$
$f(x) = \lambda e^{-\lambda x}$ for x

$P(X=i, Y=j) = P(X=i)P(Y=j)$, $\forall i,j \in R$

in continuous case

$x>0, y>0$
$e^{-x}\cdot\frac{1}{2}e^{-\frac{y}{2}} = \frac{1}{2}e^{-x-\frac{y}{2}}$

$x>0, y\leq 0$
$e^{-x}\cdot 0 = 0$

$x\leq 0, y>0$
$0\cdot\frac{1}{2}e^{-\frac{1}{2}y} = 0$.

$x\leq 0, y\leq 0$
$0\cdot 0 = 0$

$\Rightarrow X, Y$ are indep.
$\Rightarrow COV(X,Y) = 0$

$cov = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$
$corr = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

... of waiting for the next ... $\times 50 = 8.75$ ...

Ex:

X sex \ Y age group	<30	30-50	>50	
M	16	20	4	40
W	24	20	16	60
	40	40	20	100

1). Marginal distribution of X and Y.

2). Are X, Y indep?

3). Are the events

A. a customer is a man

B. — " — <30

independence?

4). Are the events

A a woman

B. >50 years

indep?

5) If the customers in b), c), d). differ. explain why?

Ex 28: 100 shops, The probability that a shop will be in black at the end of year is 0.9; Find probability that at the end of year, at the least 85 shops will be in the black.

Use CLT (Central limit theorem): $P\left(\frac{\sum_{i=1}^{n} X_i - nEX_1}{\sqrt{n\operatorname{var}X_1}} \le a\right) = \Phi(a)$.

$X_i = 1$ the i-th shop is in black at the end of year

$X_i = 0$ the — " — in the lost — " —

$X_i \sim Bern(0.9) \Rightarrow EX_i = 0.9$

$\operatorname{var}X_i = 0.9\times 0.1 = 0.09.\quad i = 1,2,\dots 100.$

$$Z = \frac{\sum_{i=1}^{100} X_i - 100\cdot 0.9}{\sqrt{100\cdot 0.09}}$$

Then $P\left(\sum_{i=1}^{100} X_i \ge 85\right) = P\left(\frac{\sum_{i=1}^{100} X_i - 100\times 0.9}{\sqrt{100\times 0.09}} \ge \frac{85 - 100\times 0.9}{\sqrt{100\times 0.09}}\right) = P(Z \ge -1.\bar{6}) = 1 - P(Z < -1.\bar{6})$

$= 1 - \Phi(-1.\bar{6}) = 1 - (1 - \Phi(1.\bar{6}))$

$= \Phi(1.\bar{6}) = 0.953$

29: 我乘8路tram去school. 电车到达之间的时间间隔为10 min. 求我在24个工作日内,在前往教职员工和回家路上,等待电车的总时间少于3小时的概率.

Use CLT: $X_i \dots$ the time of waiting for a tram during i-th Journey. $i = 1,2\dots 48$

$X_i \sim U(0,10) \Rightarrow EX_1 = 5.\quad \operatorname{Var}X_1 = \frac{25}{3}$

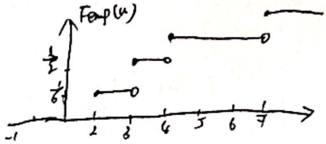
$$Z = \frac{\sum_{i=1}^{48} X_i - 48\cdot 5}{\sqrt{48\cdot\frac{25}{3}}}$$

$P\left(\sum_{i=1}^{48} X_i \le 180\right) = P\left(\frac{\sum_{i=1}^{48} X_i - 48\times 5}{\sqrt{48\times\frac{25}{3}}} \le \frac{180 - 48\times 5}{\sqrt{48\times\frac{25}{3}}}\right) = P(Z \le -3) = \Phi(-3) = 1 - \Phi(3) = 1 - 0.9\dots$

$= 0.0013.$

Theoretical distr from $F(u) = P(X \le u)\ \forall u \in \mathbb{R}$

Estimate of ↑ is $F_{emp}(u) = \frac{\# x_1, \dots x_i \le u}{n}$.



2, 3, 3, 4, 7, 7

#31: Time intervals between two breakdowns of a device ~~were~~ were 4 days, 7 days, 12 days, 2.5 days ~~and~~ a 24.5 days. The time intervals are supposed to come from exponential distribution. Use the maximum likelihood method to estimate λ.

$X \sim Exp(\lambda) \Rightarrow EX = \frac{1}{\lambda}$

$f(x) = \lambda e^{-\lambda x}$ for $x > 0$

$= 0$ for $x \le 0$

$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \frac{1}{5}(7 + 4 + \dots + 24.5) = 10$

$\frac{1}{\hat{\lambda}} = 10 \Rightarrow \hat{\lambda} = \frac{1}{10}$.

$L(\lambda) = \lambda e^{-\lambda \cdot 2.5} \cdots \lambda e^{-\lambda \cdot 24.5} = \lambda^5 (e^{-\lambda(2.5 + \dots + 24.5)})$

$l(\lambda) = \ln(\lambda^5) + \ln(e^{-\lambda(2.5 + \dots + 24.5)}) = 5\ln(\lambda) - \lambda \cdot 50$.

$\frac{\partial l(\lambda)}{\partial \lambda} = \frac{5}{\lambda} - 50 = 0. \Rightarrow \hat{\lambda} = \frac{1}{10}$.

Ex:

Pizza → ham 2 mushroom, cheese, chicken

ham 2 mushroom	cheese	chicken
50	30	20

(Customers.)

Suppose that $P(ham) = a + b$

$P(cheese) = a = \frac{1}{3}$

$P(chicken) = a - b$.

$a + b + a + a - b = 1$

$a = \frac{1}{3}$

Estimate, using one of the methods (MME or MLE), the parameter a and b.

①. $L(b) = (\frac{1}{3} + b)^{50} (\frac{1}{3})^{30} \cdot (\frac{1}{3} - b)^{20}$

②. $l(b) = \ln((\frac{1}{3} + b))^{50} + \ln(\frac{1}{3})^{30} + \ln(\frac{1}{3} - b)^{20} = 50\ln(\frac{1}{3} + b) + 30\ln(\frac{1}{3})$

③ $\frac{\partial l(b)}{\partial b} = \frac{50}{\frac{1}{3} + b} = 0 \Rightarrow \hat{b} = \frac{1}{7}$

Ex: We observe 10 times the times of waiting for a bus, where we know neither schedule nor the intervals between arrivals. We observe: 15, 3, 2, 7, 6, 2, 5.5, 4, 5.5, 1 $\le b$

What distribution do we use for time of waiting for the bus?

Estimate its parameters by MME and MLE. $X \sim U(0, b)$

$EX = \frac{b}{2} \approx \frac{1}{10} \sum_{i=1}^{n} x_i$

$\frac{b}{2} = \frac{33.5}{10} \Rightarrow \hat{b} = 6.7$

① $L(b) = f(x_1) \cdot f(x_2) \dots$ $f(x_1) = (\frac{1}{b})^{10} = \frac{1}{b^{10}}$ $\hat{b} = 7$

maximize $l(b) = \ln(\frac{1}{b^{10}}) = \ln(1) - 10\ln(b)$

... is 6. $(3,1)$ ... $(6,6)$
$(4,1)$ ... $(5,6)$
$(5,1)$ ... $(6,6)$.

Ex 32: 根据64组医生的数据计算出平均每天23名患者. 样本方差 = 36.
患者分布未知. 构建每天预测患者数的 95% CI.

$n = 64$. $\bar{X}_{64} = 23$. $S_{64}^2 = 36$. $\alpha = 0.05$, we can find $u_{0.975} = 1.96$.

$(1-\alpha)\cdot 100\%$ CI = for $EX \in (\bar{X}_n \pm U_{1-\frac{\alpha}{2}} \cdot \frac{S_n}{\sqrt{n}})$

95% CI $= (23 - \frac{\sqrt{36}}{\sqrt{64}} \times 1.96,\ 23 + 1.96 \cdot \frac{6}{8}) =$

黑板:
$EX \in (\bar{X}_{64} \pm U_{0.975} \cdot \frac{S_{64}}{\sqrt{n}}$

$EX \in (23 \pm 1.96 \times \frac{\sqrt{36}}{\sqrt{64}})$

$EX \in (23 \pm 1.5)$

$EX \in (21.5, 24.5)$

EX 30(1). 数据: $\boxed{\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad . \quad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$

$n = 9$. $\bar{X} = \frac{0+2+\dots}{9} = 3$. $\alpha = 0.01$. 99% CI. $1 - \frac{\alpha}{2} = 1 - \frac{0.01}{2} = 0.995$

$S_9^2 = \frac{1}{8}[(0-3)^2 + (2-3)^2 + \dots\dots] = 4.43$. $u_{0.995} = 2.58$

$EX \in (3 \pm 2.58 \times \frac{\sqrt{4.43}}{\sqrt{9}})$

Data $X \sim P_o(\lambda)$

$EX = \operatorname{var} X = \lambda$

$\underline{\underline{\bar{X}_n}}$ 无 S_n^2

$EX \in (3 \pm 2.58 \cdot \frac{\sqrt{3}}{\sqrt{9}})$

EX: Last week, we sample a ball from a box with white and black balls, we sampled 12 white and 8 black balls. Find 95% CI for the probability that we sample a white ball. (use an analoge approach as

$X_i = 1$, if the i-th sampled ball is white

$X = 0$, if —— " —— black.

$\frac{3}{5} + \frac{2}{5}$

$X_i \sim alt(p)$

$EX = p$

$\operatorname{var} X = p(1-p)$

$n = 20$.

$\bar{X}_{20} = \frac{12}{20} = \frac{3}{5}$ → $p(1-p) = \frac{3}{5} \times \frac{2}{5}$

$S_{20}^2 = \frac{6}{25} = \frac{96}{400}$.

$EX \in (\frac{3}{5} \pm 1.96 \cdot \frac{\sqrt{\frac{6}{25}}}{\sqrt{20}})$ $\bar{X}_{20}(1 - \bar{X}_{20})$.

$EX \in (\frac{3}{5} \pm 1.96 \cdot \sqrt{\frac{6}{500}})$

$$S_{20}^2 = \frac{1}{19}\Big[\overbrace{(1-\tfrac{3}{5})^2 + (1-\tfrac{3}{5})^2 + \dots}^{12\times} + \underbrace{(0-\tfrac{3}{5})^2 + \dots + (0-\tfrac{3}{5})^2}_{8\times}\Big] = \frac{1}{19}\Big[12 \cdot \frac{4}{25} + 8 \cdot \frac{9}{25}\Big] =$$