Assignment 3

Due date : 5/6/2022

1. (Slightly modified version of Ex 4-35 on page 162) Suppose the diameter at breast height (in.) of trees of a certain type is normally distributed with $\mu = 9$ and $\sigma = 2$.

a. What is the probability that the diameter of a randomly selected tree will be at least 11 in.? $\underline{0.1587}$

b. What is the probability that the diameter of a randomly selected tree will be between 7 and 13 in.? $\underline{0.6827}$

c. What value c is such that the interval (9-c, 9+c) includes 98% of all diameter values? $\underline{2.33}$

d. If six trees are independently selected, what is the probability that at least two has a diameter exceeding 11 in.? $\underline{0.2441}$

a. $P(X \geq 11) = ?$ $\quad z = \frac{11-9}{2}$, $\quad P(Z \geq 1) = 0.1587.$

b. $P(7 \leq X \leq 11) = ?$ $\quad z = \frac{7-9}{2} = -1$, $\quad z = \frac{11-9}{2} = 1$ $\quad P(-1 \leq Z \leq 1) = 0.6827$

c. $P(9-c \leq X \leq 9+c) \rightarrow P(-c \leq Z \leq c) = 98\% = 0.98$

$P(Z \geq c) = 0.01$ $\qquad c \cong 2.33$

d. $X \rightarrow$ event exceeding 11 inch

$P(X) = 0.1587 \quad (\because a)$
$= p$

at least 2 $\rightarrow$ Total $- 0 - 1$ $= 1 - {}_6C_0\, p^0 (1-p)^6 - {}_6C_1\, p^1 (1-p)^5$

$= 0.2441$

2. (Slightly modified version of Ex 4-47 on page 164) Suppose only 75% of all drivers in a certain state regularly wear a seat belt. A random sample of 500 drivers is selected and let X denote the number of drivers who wear a seat belt.

a) Calculate the E(X) and V(X) $\underline{E(X) = np = 375}$ $\underline{V(X) = 93.75 = np(1-p)}$

b) Calculate the probability that X is between 360 and 400 (inclusive)($P(360 \leq X \leq 400)$) $\underline{0.9280}$

c) Calculate the probability that X is between 360 and 400 inclusive by approximation. $\underline{0.9287}$

a. $p = 0.75 \quad n = 500.$

$E(X) = np \quad V(X) = np(1-p) \quad \begin{cases} \mu = 375 \\ \sigma = 9.6825 \end{cases}$

b. $\sum_{x=360}^{400} {}_{500}C_x \, p^x (1-p)^{500-x}$ = pbinom(400, 500, 0.75) − pbinom(360, 520, 0.75)

$= 0.9280$

c. $\Phi\left(\frac{400+0.5-\mu}{\sigma}\right) - \Phi\left(\frac{360+0.5-\mu}{\sigma}\right)$

$= \Phi(2.6336) - \Phi(-1.4975) = 0.9958 - 0.0671$

$= 0.9287$

pnorm(400.5, 375, 9.6825) − pnorm(360.5, 375, 9.6825)

$= 0.9287.$

3. (Slightly modified version of Ex 4-59 on page 170) Let $X =$ the time between two successive arrivals at the drive-up window of a local bank. If X has an exponential distribution with $\lambda = 3$, compute the following:

a. The expected time between two successive arrivals $E(X) = \frac{1}{\lambda} = \frac{1}{3}$

b. The standard deviation of the time between successive arrivals $V(X) = \frac{1}{\lambda^2} = \frac{1}{9}$ $\sigma = \sqrt{\frac{1}{9}} = \frac{1}{3}$

c. Calculate the median $\tilde{\mu}$. $1 - e^{-3\lambda} = 0.5$, $e^{-3\lambda} = 0.5$, $\lambda = 0.231$

d. Calculate $P(X \leq 2)$ $1 - e^{-3 \cdot 2} = 0.9975$

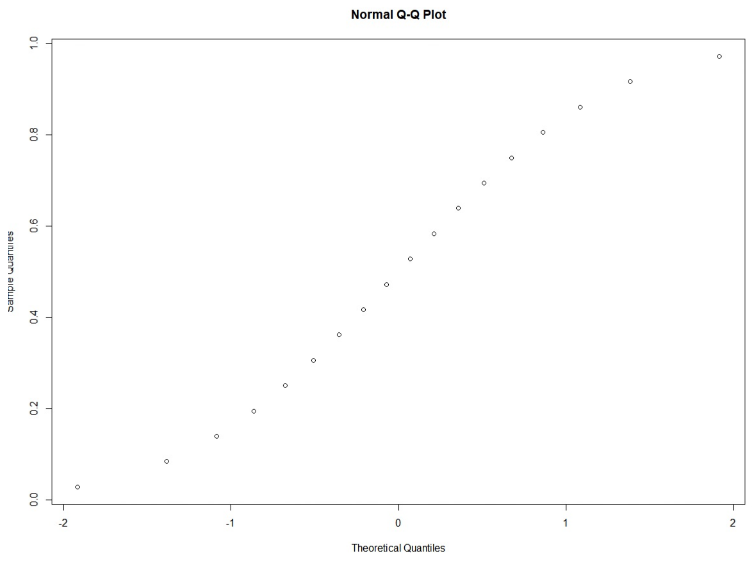
b. Calculate $P(1 \leq X \leq 2)$. $(1-e^{-6}) - (1-e^{-3}) = e^{-3} - e^{-6} = 0.0473$

4. (Ex 4-90 on page 187) Consider the following sample of size n=18 observations on toughness for high-strength concrete; values of $p_i = \frac{\mathrm{i}-0.5}{18}$ are also given. Construct a normal and Weibull probability plot and comment.

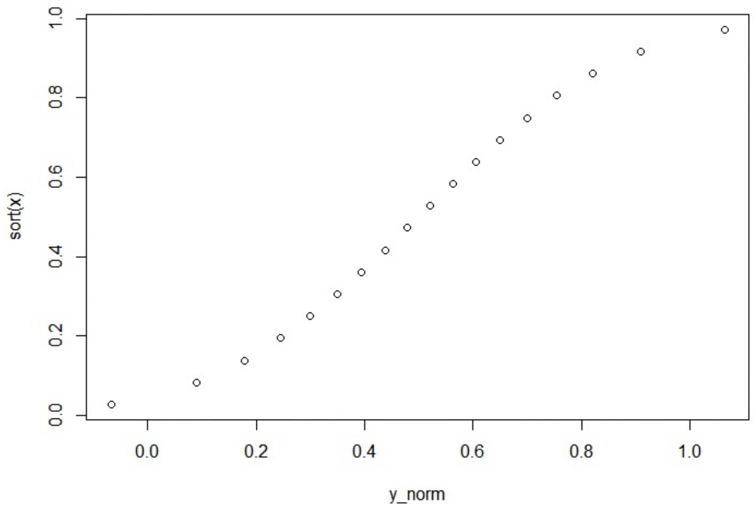
Observation	.47	.58	.65	.69	.72	.74
	.0278	.0833	.1389	.1944	.2500	.3056
Observation	.77	.79	.80	.81	.82	.84
	.3611	.4167	.4722	.5278	.5833	.6389
Observation	.86	.89	.91	.95	1.01	1.04
	.6944	.7500	.8056	.8611	.9167	.9722

Cf) The normal probability plot and Weibull probability plot of Ex 4-92 can be obtained as follows:

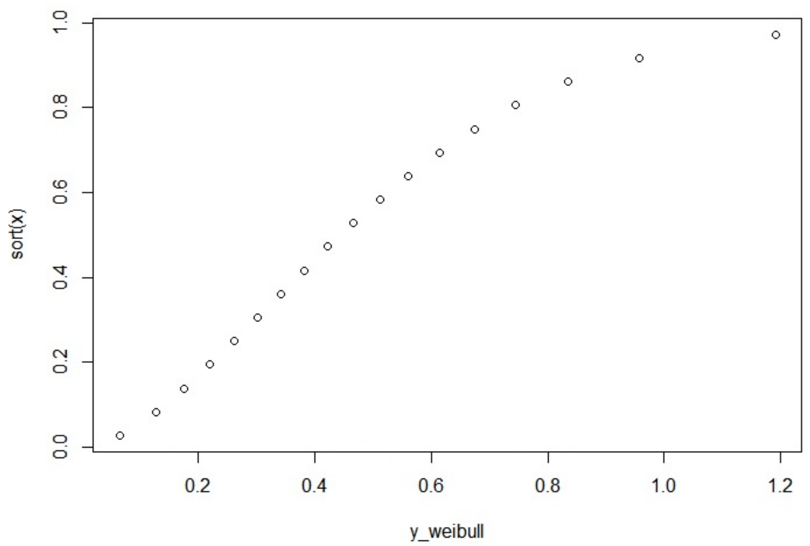
```
x <- c(47.1, 68.1, 68.1, 90.8, 103.6, 106.0, 115.0, 126.0, 146.6, 229.0, 240.0,
240.0, 278.0, 278.0, 289.0, 289.0, 367.0, 385.9, 392.0, 505.0)
sort(x)

# normal distribution Q-Q plot
qqnorm(x)

# normal distribution Q-Q plot
loc <- seq(0.025, 0.975, by=0.05) #p_i = (i-0.5)/n = (i-0.5)/20
y_norm <- qnorm(loc, mean(x), sd(x))
plot(y_norm, sort(x))

# weibull distribution Q-Q plot
library(fitdistrplus)
fit_w <- fitdist(x, "weibull")
fit_w$estimate[1] #estimate the parameters of weibull distribution
y_weibull <- qweibull(loc, fit_w$estimate[1], fit_w$estimate[2])
plot(y_weibull, sort(x))
```



Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
0.0
0.2
0.4
0.6
0.8
1.0
-2
-1
0
1
2



sort(x)
y_norm
0.0
0.2
0.4
0.6
0.8
1.0



sort(x)
y_weibull
0.0
0.2
0.4
0.6
0.8
1.0
1.2

5. (Ex 5-31 on page 212) Annie and Alvie have agreed to meet for lunch between noon (0:00 P.M.) and 1:00 P.M. Denote Annie's arrival time by X, Alvie's by Y, and suppose X and Y are independent with pdf's

$$f(x) = \begin{cases} 3x^2 & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$$

$$f(y) = \begin{cases} 2y & 0 \le y \le 1 \\ 0 & otherwise \end{cases}$$

What is the expected amount of time that the one who arrives first must wait for the other person? [Hint: h(X, Y)=| X-Y |.]

$$f_{XY}(x,y) = 6x^2y \quad (0 \le x \le 1) \quad (0 \le y \le 1)$$

$$h(x,y) = |x-y|$$

$$\int_0^1 \int_0^1 |x-y|\, 6x^2y \, dy\, dx = \int_0^1 \left\{ \int_0^x (x-y)\, 6x^2y \, dy \right\} dx + \int_0^1 \left\{ \int_0^y (y-x)\, 6x^2y \, dy \right\} dx$$

$$= \int_0^1 \left\{ \int_0^x 6x^3y - 6x^2y^2 \, dy \right\} dx + \int_0^1 \left\{ \int_0^y 6x^2y^2 - 6x^3y \, dx \right\} dy$$

$$= \int_0^1 \left[3x^3y^2 - 2x^2y^3 \right]_0^x dx + \int_0^1 \left[2x^3y^2 - \frac{3}{2}x^4y \right]_0^y dy$$

$$= \int_0^1 (3x^5 - 2x^5)\, dx + \int_0^1 \left(2y^5 - \frac{3}{2}y^5\right) dy = \int_0^1 x^5\, dx + \int_0^1 \frac{1}{2}y^5\, dy = \left[\frac{1}{6}x^6\right]_0^1 + \left[\frac{1}{12}y^6\right]_0^1$$

$$= \frac{1}{6} + \frac{1}{12} = \frac{3}{12} = \boxed{\frac{1}{4}}$$

6. (Ex 5-77 on page 236) A health-food store stocks two different brands of a certain type of grain. Let X = the amount (lb) of brand A on hand and Y = the amount of brand B on hand. Suppose the joint pdf of X and Y is

$$\mathrm{f(x, y)} = \begin{cases} kxy & x \geq 0, 20 \leq x + y \leq 30 \\ 0 & otherwise \end{cases}$$

a) Draw the region of positive density and determine the value of k. $k = 0.000036923076$

b) Are X and Y independent? Answer by first deriving the marginal pdf of each variable. No.

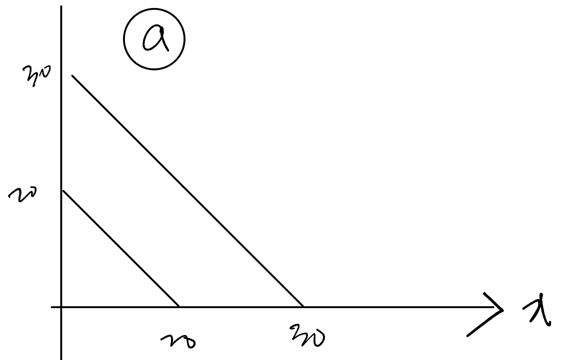
c) Compute P(X+Y≤25). 0.3578076844

d) What is the expected total amount of this grain on hand? $\boxed{25.96923012}$.

e) Compute Cov(X, Y) and Corr(X, Y). $\boxed{-32.1899753}$ $\boxed{-0.8937898132}$

f) What is the variance of the total amount of grain on hand?

a



$$\int_0^{20}\int_{20-x}^{30-x} kxy \, dy\,dx + \int_{20}^{30}\int_0^{30-x} kxy \, dy\,dx = 1$$

$$= \int_0^{20} \left[kx \times \frac{1}{2}y^2\right]_{20-x}^{30-x} dx + \int_{20}^{30} \left[kx \times \frac{1}{2}y^2\right]_0^{30-x} dx$$

$$= \int_0^{20} \frac{kx}{2} \times \left\{(30-x)^2 - (20-x)^2\right\} dx + \int_{20}^{30} \frac{kx}{2}(30-x)^2 dx$$

$$= \int_0^{30} \frac{kx}{2}(30-x)^2 - \int_0^{20} (20-x)^2 \times \frac{kx}{2} dx = 1 \quad = \frac{k}{2}\left\{\int_0^{30} x(30-x)^2 dx - \int_0^{20} x(20-x)^2 dx\right\}$$

$$= \frac{k}{2}\left\{67500 - \frac{40000}{3}\right\} = 1, \quad k = 0.000036923076$$

(b) if X & Y are independent $\therefore f(X)\cdot f(Y) = f(X,Y)$

$$f(X) = \int_{20-x}^{30-x} kxy\,dy = \left[\frac{k}{2}xy^2\right]_{20-x}^{30-x}$$

$$= \frac{k}{2}x\left\{(30-x)^2-(20-x)^2\right\} = \frac{k}{2}x\left\{(30-x+20-x)(30-x-20+x)\right\}$$

$$= \frac{k}{2}x\left\{(50-2x)\times 10\right\} = 5kx(50-2x) = 10kx(25-x) \quad (0\le x\le 20)$$

$$\int_{0}^{30-x} kxy\,dy = \left[\frac{1}{2}kxy^2\right]_0^{30-x} = \frac{1}{2}kx(30-x)^2 \quad (20\le x\le 30)$$

$$f(X)\begin{cases} kx(250-10x) & (0\le x\le 20) \\ \frac{1}{2}kx(30-x)^2 & (20\le x\le 30)\end{cases}$$

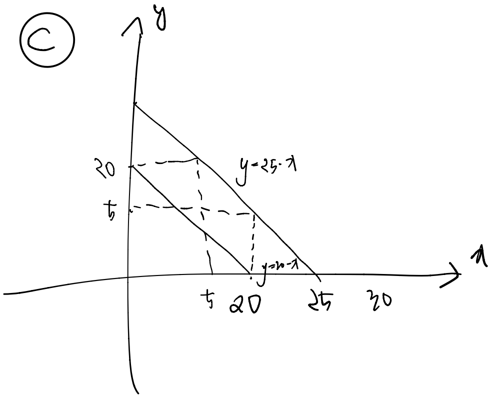
$$f(y) = \int_{20-y}^{30-y} kxy\,dx = ky(250y-10y^2) \quad (0\le y\le 20)$$

$$= \int_{0}^{30-y} kxy\,dx = \frac{1}{2}ky(30-y)^2 \quad (20\le y\le 30)$$

$$f(y)\begin{cases} ky(250y-10y^2) & (0\le y\le 20) \\ \frac{1}{2}ky(30-y)^2 & (20\le y\le 30)\end{cases}$$

$$\boxed{f(X)\times f(Y) \ne f(X,Y) \quad \therefore \text{Not independent}}$$

ⓒ



$$\int_0^{20}\int_{20-x}^{25-x} kxy\,dy\,dx + \int_{20}^{25}\int_0^{25-x} kxy\,dy\,dx.$$

$$= k\int_0^{20}\left[\frac{1}{2}xy^2\right]_{20-x}^{25-x} dx + k\int_{20}^{25}\left[\frac{1}{2}xy^2\right]_0^{25-x} dx$$

$$= k\int_0^{20}\frac{1}{2}x\left\{(45-2x)(5)\right\}dx + k\int_{20}^{25}\frac{1}{2}x(25-x)^2\,dx = k\left(9166.666667 + 442.7083333\right)$$

$$= 0.3548076824$$

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

$$\int_0^{20} x f(x)dx + \int_{20}^{30} x f(x)dx + \int_0^{20} y f(y)dy + \int_{20}^{30} y f(y)dy$$

$$= \left\{ \int_0^{20} xk(250x - 10x^2)dx + \int_{20}^{30} \frac{kx}{2}(30-x)^2 dx \right\} \times 2$$

$$= k \times 2 \times \left\{ 266666.6667 + 85000 \right\} = \boxed{25.96923012}$$

$$E(X) = E(Y) = 12.98461506$$

(e)

$$E(XY) - E(X)E(Y)$$

$$= E(XY) - 168.6002283$$

$$= 136.410253 - 168.6002283$$

$$= -32.1899753$$

$$E(XY) = \int_0^{20}\int_{20-x}^{30-x} kx^2y^2\,dy\,dx + \int_{20}^{30}\int_0^{30-x} kx^2y^2\,dy\,dx$$

$$= \int_0^{20}\left[\frac{k}{3}x^2y^3\right]_{20-x}^{30-x} dx + \int_{20}^{30}\left[\frac{k}{3}x^2y^3\right]_0^{30-x} dx$$

$$= k\int_0^{20} \frac{1}{3}x^2\left\{(30-x)^3-(20-x)^3\right\}dx + k\int_{20}^{30}\frac{1}{3}x^2\cdot(30-x)^3\,dx$$

$$= k\left\{3888888.889 + 405555.556\right\} = 136.810253$$

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_x\sigma_y}$$ / $$\sigma_x = \sqrt{E(X^2) - E(X)^2}$$

$$E(X^2) = k\left\{\int_0^{20} x^2\cdot x(250-10x)\,dx + \int_{20}^{30} x^2\cdot\frac{x}{2}(30-x)^2\,dx\right\}$$

$$= k\left\{3600000 + 1941666.667\right\} = 204.6153795$$

$$\sigma_x^2 = 204.6153795 - 12.98861506^2$$

$$= 36.0151524$$

$$\sigma_x = \sigma_y$$

$$\mathrm{Corr}(X,Y) = \frac{-32.1899753}{36.0151524} = -0.893789813$$

$$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X,Y)$$
$$= 7.65035688$$

7. (Ex 5-73 on page 235) Suppose the expected tensile strength of type-A steel is 105 ksi and the standard deviation of tensile strength is 8 ksi. For type-B steel, suppose the expected tensile strength and standard deviation of tensile strength are 100 ksi and 6 ksi, respectively. Let $\bar{X}$ = the sample average tensile strength of a random sample of 40 type-A specimens, and let $\bar{Y}$ = the sample average tensile strength of a random sample of 35 type-B specimens.

a) What is the approximate distribution of $\bar{X}$? Of $\bar{Y}$? $\bar{X} \sim N(105, \sqrt{\frac{8}{40}}^2)$ $\bar{Y} \sim N(100, \sqrt{\frac{6}{35}}^2)$

b) What is the approximate distribution of $\bar{X}-\bar{Y}$? Justify your answer. $\bar{X}-\bar{Y} \sim N(5, \sqrt{\frac{9^2}{35}}^2)$

c) Calculate (approximately) $P(-1 \le \bar{X}-\bar{Y} \le 1)$. 0.0067

$\bar{X} =$ $\mu = 105$, $\sigma = 8$, $n = 40$

$\bar{Y} =$ $\mu = 100$, $\sigma = 6$, $n = 35$

$\bar{X} \sim N\left(105, \left(\frac{8}{\sqrt{40}}\right)^2\right)$

$\bar{Y} \sim N\left(100, \left(\frac{6}{\sqrt{35}}\right)^2\right)$

b)

$Y = X_1 - X_2$

$E(Y) = E(X_1) - E(X_2)$

$V(Y) = V(X_1 - X_2) = 1^2 V(X_1) + (-1)^2 V(X_2)$

$= V(X_1) + V(X_2) = \sigma_1^2 + \sigma_2^2$

$\sigma_Y^2 = \sigma_1^2 + \sigma_2^2$ $\sigma_Y = \sqrt{\sigma_1^2 + \sigma_2^2}$

$$\boxed{\begin{aligned}\bar{X}-\bar{Y} &\sim N\left(105-100, \sqrt{\frac{8}{5}+\frac{36}{35}}^{\,2}\right)\\ &= N\left(5, \sqrt{\frac{92}{35}}^{\,2}\right)\end{aligned}}$$

C) $P(-1 \le \bar{X}-\bar{Y} \le 1)$

$$= P\left(\frac{-1-5}{\sqrt{\frac{92}{35}}} \le Z \le \frac{1-5}{\sqrt{\frac{92}{35}}}\right)$$

$$= P(-3.7 \le Z \le -2.4671) = 0.0067$$