Valeria Duran

Math 6358

Homework 3

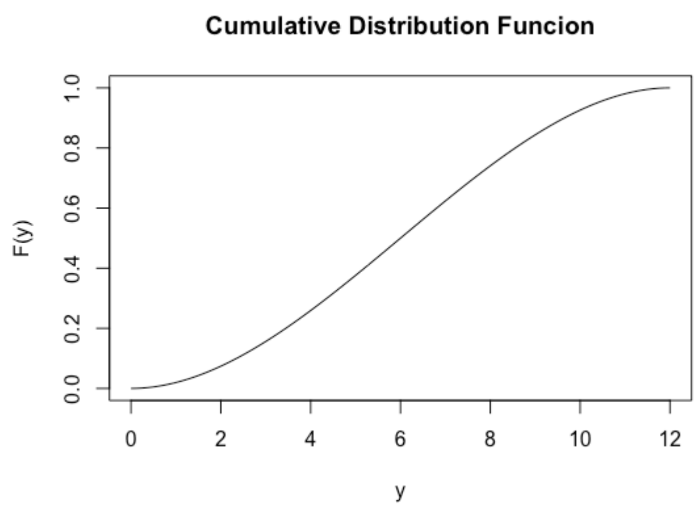
Problem 1.

A. The CDF is:

F(y) = { 0 for y <0

$\frac{1}{48}(y^2 - \frac{y^3}{18})$ for $0 \le y \le 12$

1 for y > 12}



B. P(Y ≤ 4) =

```
f <- function(y) {(y/24)*(1-(y/12))}
integrate(f, lower = 0, upper = 4)
```

0.2592593

P(Y>6) =

```
> integrate(f, lower = 6, upper = 12)
```

0.5

P(4 ≤ Y ≤ 6) =

```
> integrate(f, lower = 4, upper = 6)
```

0.2407407

C. E(Y)=

> e <- function(y) {(y/24)*(1-(y/12))*y}

> integrate(e, lower = 0, upper = 12)

6

$E(Y^2)$ =

> e <- function(y) {(y/24)*(1-(y/12))*y^2}

> integrate(e, lower = 0, upper = 12)

43.2

Var(Y) = $E(Y^2)$ - $E(Y)^2$

> 43.2 - 6^2

[1] **7.2**

D. P(|Y - μ| > 2)

= 1 - P(-2 < Y - 6 < 2)

= 1 - P(4 < Y < 8)

= > f <- function(y) {(y/24)*(1-(y/12))}

> integrate(f, lower = 4, upper = 8)

0.4814815 with absolute error < 5.3e-15

> 1 - 0.4814815

[1] **0.5185185**

E.

Y = bar length

Y - 12 = smaller bar length

Since E(Y) = 6;

0 < Y < 6 : y < 12 - Y —> y

Y = 6 : y = 12 - Y

6 < Y < 12 : y > 12 - Y —> y = 12 - Y

> d <- function(y) {(y/24)*(1-(y/12))*y}

> g <- function(y) {(y/24)*(1-(y/12))*(12-y)}

> integrate(d, lower = 0, upper = 6)

1.875 with absolute error < 2.1e-14

> integrate(g, lower = 6, upper = 12)

1.875 with absolute error < 2.1e-14

1.875 + 1.875 = **3.75**

Problem 2.

Problem 3.

Normal distribution with mean = 137.2 and sd = 1.6. stated content is 135

A. > 1 - pnorm(135, 137.2, 1.6)

[1] **0.9154343**

B. P(X ≥ 8)

binomial distribution n=10; p = 0.9154

> 1 - pbinom(7, 10, 0.9154)

[1] **0.9537763**

C. > qnorm(.05)

[1] -1.644854

$135 = 137.2 - 1.644854\sigma$

$1.644854\sigma = 137.2 - 135$

> (137.2 - 135)/ 1.644854

[1] **1.337505**

Problem 4.

A. $E[X] = e^{\mu+\sigma^2/2}$

E[X] = 68.03

$Var(X) = e^{2\mu+\sigma^2} * (e^{\sigma^2} - 1)$

Var(X) = 14907.17

SD(X) = 122.09

B. $P(50 \leq X \leq 250) =$

> plnorm(250, 3.5, 1.2) - plnorm(50, 3.5, 1.2)

[1] **0.319629**

C. plnorm(68.03, 3.5, 1.2)

[1] **0.7257327**

The probability is not 0.5 because the lognormal distribution is not symmetric.

Problem 5.

A. Marginal pmf of X:

For X = 20: 0.05 + 0.05 + 0.1

[1] **0.2**

For X = 25: 0.05 + 0.10 + 0.35

[1] **0.5**

For X = 30: 0 + 0.20 + 0.10

[1] **0.3**

Marginal pmf for Y:

For Y = 20: 0.05 + 0.05 + 0

[1] **0.1**

For Y = 25: 0.05 + 0.10 + 0.20

[1] **0.35**

For Y = 30: 0.10 + 0.35 + 0.10

[1] **0.55**

B. $P(X \leq 25 \text{ and } Y \leq 25) = P(X = 20 \text{ and } Y = 20) + P(X = 20 \text{ and } Y = 25) + P(X = 25 \text{ and } Y = 20) + P(X = 25 \text{ and } Y = 25) = 0.05 + 0.05 + 0.05 + 0.10$

[1] **0.25**

C. For it to be independent, the probability distribution of X has to equal the conditional probability of P(X| Y=20):

> 0.05/0.1 #For X=20

[1] 0.5

> 0.05/0.1 #For X=25

[1] 0.5

> 0/0.1 #For X=30

[1] 0

Since the conditional probability and the probability distribution of X are not the same, **X and Y are not independent.**

D. E[X] =

> 20*0.2 + 25*0.5 + 30*0.3 #expected value of X

[1] 25.5

E[Y] =

> 20*0.1 + 25*0.35 + 30*0.55 #expected value of Y

[1] 27.25

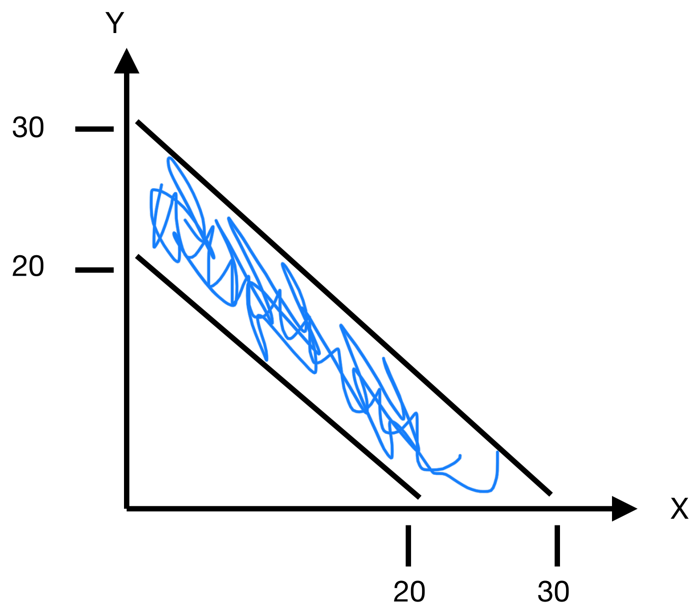
E[X and Y] =

> 25.5 + 27.25 #expected value of X and Y

[1] **52.75**

E. E[Y-X] = E[Y] - E[X] = 27.25 - 25.5 = **1.75**

Problem 6.



$$\int_0^{30}\int_0^{30-x} f(x,y)\,dy\,dx - \int_0^{20}\int_0^{20-x} f(x,y)\,dy\,dx = 1$$

Problem 7.

n	1	2	3	4
p(n)	0.4	0.3	0.2	0.1

A. E(Y| X= x) = x*n*p(n) = x*(1*0.4 + 2*0.3 + 3*0.2 + 4*0.1)

2x

Var(Y| X = x) = E(Y^2 | X = x) - E(Y| X = x)^2 =

E(Y^2 | X = x) = x^2*n^2*p(n) = > (1^2)*0.4 + (2^2)*0.3 + (3^2)*0.2 + (4^2)*0.1

[1] 5

Var(Y| X = x) = $5x^2 - 4x^2 = x^2$

B. E[Y] = E[E(Y| X= x)] = E[2x] = 2E[X] = 2×20 **= 40**

C. Var[Y] = V[E(Y| X= x)] + E[V(Y| X= x)] =

V[E(Y| X= x)] = V[2x] = 4V[x] = 80

E[V(Y| X= x)] = E[X] = 20

Var[Y] = 80 + 20 = **100**

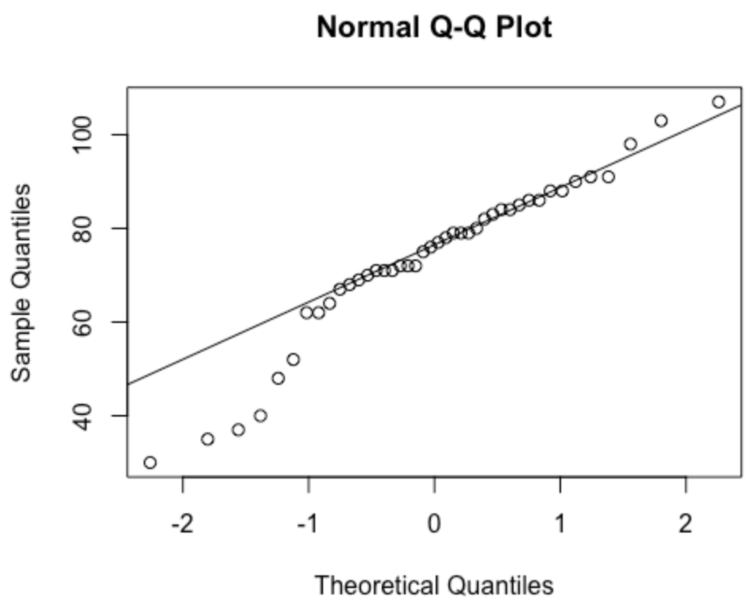
Problem 8.

> airpollution <- data.frame(airpollution)

> attach(airpollution)

A. > qqnorm(Solar)

> qqline(Solar)



B. > shapiro.test(Solar)

Shapiro-Wilk normality test

data: Solar

W = 0.93883, p-value = 0.02601

The W value is large and the p-value is less than 0.05, indicating that the solar radiation measurements are normally distributed. The QQ plot in part a shows possible outliers in the data;

apart from this, the data still seems to be normally distributed based off of the plot. This shows that the Shapiro-walker test in part b gives the same results as in part a.

Problem 9.

Per day (X):

Mean = 13 oz

SD = 2 oz

14 days (W):

E[W] = 14E[X] = 14 * 13

[1] 182

SD(W) = sqrt(14) * 2

[1] 7.483315

2 6-packs of 16 oz sodas is 192 oz in total

> pnorm(192, 182, 7.483315)

[1] **0.9092754**

We should worry about the validity of the normality assumption here since the sample size is not large enough (14 days) and might not actually provide a normal distribution.

Problem 10.

E[X] = 70

Var[X] = 9

E[Y] = 170

Var[Y] = 400

$\rho = 0.9$

A.

$$(X,Y) \sim BVN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$$

$$(X,Y) \sim BVN(70, 170, 3, 20, 0.9)$$

$$(Y|X=x) \sim N\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}\,(x-\mu_x), \sigma_y^2(1-\rho^2)\right)$$

$$(Y|X=x) \sim N(170 + 0.9 * \frac{20}{3}\,(x-70), 400(1-0.9^2))$$

$$(Y|X=x) \sim N(170 + \ 6(x-70), 76)$$

$$(Y|X=68) \sim N(170 + 6(68-70), 76)$$

$$(Y|X=68) \sim N(158, 76)$$

B.

$$(Y|X=x) \sim N(170 + \ 6(x-70), 76)$$

$$(Y|X=70) \sim N(170 + \ 6(70-70), 76)$$

$$(Y|X=70) \sim N(170, 76)$$

Both distributions of Y at X = 68 and X = 70 are univariate normal. However, the univariate distribution of Y at X = 68 has a mean of 158 and a variance of 76 while the distribution of Y at X = 70 has a mean of 170 and a variance of 76.

C.

$$(Y|X=x) \sim N(170 + \ 6(x-70), 76)$$

$$(Y|X=72) \sim N(170 + \ 6(72-70), 76)$$

$$(Y|X=72) \sim N(182, 76)$$

$$P(Y < 180\ |X=72) = P\left[\frac{Y-\mu}{\sigma} < \frac{180-\mu}{\sigma}\right] = P\left[Z < \frac{180-182}{\sqrt{76}}\right] = P(Z < -0.2294157)$$

```
> pnorm(180, 182, sqrt(76))
[1] 0.4092729
```