INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Subject: App of IT-Basics of R

Chapter: Unit 1

Category: Practice Question Solution

Hypothesis Tests:

##Answer1

```
sd1 = 8.4
n1 = 10
x1 = 162.8
```

##a]

```
##H0 : mu = 165
##Ha : mu < 165

mu1a = 165
Y1a <- c(rnorm(n1,x1,sd1))
t.test(Y1a,mu = mu1a,alternative = "less")
##Since, p-value > 0.05, we fail to reject our null hypothesis
```

##b]

```
##H0 : mu = 158
##Ha : mu != 158
mu1b = 158
Y1b <- c(rnorm(n1,x1,sd1))
t.test(Y1b,mu = mu1b)
##Since, p-bvalue > 0.05, we fail to reject our null hypothesis
```

##Answer2

```
sd2 = 12.1
n2 = 6
mu2 = 64.2


data <- c(64.30,68.93,74.91,52.71,59.70,73.60)

##b] Hypothesis Test

##H0: Mean weight of women in the US is 64.2kg
##Ha: Mean weight of women in the US is not 64.2kg
t.test(data,mu = mu2)
##Since, p-value > 0.05, we fail to reject our null hypothesis
##The mean weight of women in the US is 64.2kg
```

##Answer3

```
n3 = 10
u3 = 22
##a] Storing values in variable rain
rain <- c(17.2,28.1,25.3,26.2,30.7,19.2,23.4,27.5,29.5,31.6)

##b] Hypothesis Test
##H0: Mean annual rainfall is equal to 22cm
##Ha: Mean annual rainfall has increased from 22cm
t.test(rain,mu = u3,alternative = "greater")
##p-value = 0.01495
##Since, p-value < 0.05, we reject our null hypothesis.
##The average rainfall has increased from it's longer value of 22cm.
```

##Answer4

```
control <- c(27.0,32.2,30.4,28.0,26.5,25.5,29.6,27.2)
additive <- c(31.4,29.9,33.2,34.4,32.0,28.7,26.1,30.3)
var(control)
var(additive)
##Output
##> var(control)
##[1] 5.054286
##> var(additive)
##[1] 6.865714
##Thus, the variances are not equal, i.e. they are not homogeneous.

##COnducting t-test
t.test(control,additive,var.equal = FALSE)
##SInce,p-value > 0.05, we fail to reject our null hypothesis.
```

##Answer5

```
n5a = 50
n5b = 40
x5a = 132
x5b = 128
sd5a = 20
sd5b = 25

Y5a <- c(rnorm(n5a,x5a,sd5a))
Y5b <- c(rnorm(n5b,x5b,sd5b))
var(Y5a)
var(Y5b)
##variances are not equal
##H0: Mean(IQs of University A) = Mean(IQs of University B)
##H1: Mean(IQs of University A) > Mean(IQs of University B)
t.test(Y5a,Y5b,var.equal = FALSE,alternative = "greater")
##Since,p-value > 0.05, we fail to reject  our null hypothesis.
```

PROBABILITY

1
##a]

##Let A be the event that the first card was red and B be the event that the
##second card is a face card
##n(A) = 26
##n(B) = 12 (Jack,Queen and King)
##Let PA be P(A), PB be P(B) and PAB be P(AnB)
PA = 26/52
PB = 12/52
PAB = 6/52
##P(B|A) = P(AnB)/P(A)
conPBA = PAB/PA
conPBA
##Answer: 0.2307692

##b]

##Let C be the event that the first card was face card and D be the event that
##the second card is an ace card.
##Using similar kind of notations
PC = 12/52
PD = 4/52
##Since, both events are independent P(CnD)=P(C)*P(D)
PCD = PC * PD
conPDC = PCD/PC
conPDC
##Answer: 0.07692308

##c]

##Let E be the event that the first card was red ace and F be the event that
##the second card is a black jack
##using similar notations,
PE = 2/52
PF = 2/52
##Since, both events are independent,
PEF = PE * PF
conPFE = PEF/PE
conPFE
##Answer: 0.03846154

##Answer2:

##Let PA, PB, PC and PX denote the probabilities of the respective events.
PA = 0.35
PB = 0.45
PC = 0.2
conXA = 0.8
conXB = 0.65
conXC = 0.3
PX = conXA*PA+conXB*PB+conXC*PC
##Let conAX, conBX, and conCX be P(A|X), P(B|X) and P(C|X) respectively.
conAX = (conXA*PA)/PX
conAX
conBX = (conXB*PB)/PX
conBX
conCX = (conXC*PC)/PX
conCX
##Answer: P(A|X) = 0.4426877, P(B|X) = 0.4624506, P(C|X) = 0.09486166

##Answer3:

##Let H be the event of getting a heart attack.
##Let P(H) = PH
##Let DA and DB be the events of giving Drug A and Drug B respectively
PDA = 50/200
PDB = 50/200
PDADB = 100/200
PH = 0.80
conHDA = (1-0.35)*PH ##P(H|DA)
conHDB = (1-0.20)*PH ##P(H|DB)
conHDADB = (1-0.35)*(1-0.20)*PH ##P(H|DA n DB)
##To find: P(DA n DB|H)
conDADBH = (conHDADB*PDADB)/(conHDA*PDA+conHDB*PDB+conHDADB*PDADB)
conDADBH
##Answer: 0.4176707

##Answer4:
##a)

a4 = ppois(2,4.2)
a4
##P(X<=2) = 0.210238

```
##b)

##P(X>=5) = 1-P(X<5) or 1-P(X<=4)
b4 = 1-ppois(4,4.2)
b4
##P(X>=5) = 0.410173

##c)

c4 = dpois(8,4.2)
c4
##P(X=8) = 0.03601115
```

##Answer5:

```
n_binomial = 30
p_binomial = 0.04

##a)

a5 = dbinom(25,n_binomial,p_binomial)
a5
##P(r=25) = 1.308245e-30

##b)

b5 = dbinom(3,n_binomial,p_binomial)
b5
##P(r=3) = 0.08630368

##c)

c5 = dbinom(5,n_binomial,p_binomial)
c5
##P(r=5) = 0.00525913
```

CHI-SQUARE TESTS

##Answer1:

```
##Loading Dataset
data_1 <- read.csv(file.choose())
data_1

##EDA
head(data_1)
print(str(data_1))
View(data_1)

##Chi-square test

##Considering NE and NW

##H0: Mean(NE) = Mean(NW)
##Ha: Mean(NE) != Mean(NW)
chisq.test(data_1$NE,data_1$NW)
##Since, p-value > 0.05, we accept our null hypothesis.

##Considering NE and SE

##H0: Mean(NE) = Mean(SE)
##Ha: Mean(NE) != Mean(SE)
chisq.test(data_1$NE,data_1$SE)
##Since, p-value > 0.05, we accept our null hypothesis.

##Considering NE and SW

##H0: Mean(NE) = Mean(SW)
##Ha: Mean(NE) != Mean(SW)
chisq.test(data_1$NE,data_1$SW)
##Since, p-value > 0.05, we accept our null hypothesis.

##Considering NW and SE

##H0: Mean(NW) = Mean(SE)
##Ha: Mean(NW) != Mean(SE)
chisq.test(data_1$NW,data_1$SE)
##Since, p-value > 0.05, we accept our null hypothesis.

##Considering NW and SW
```

```
##H0: Mean(NW) = Mean(SW)
##Ha: Mean(NW) != Mean(SW)
chisq.test(data_1$NW,data_1$SW)
##Since, p-value > 0.05, we accept our null hypothesis.

##Considering SE and SW

##H0: Mean(SE) = Mean(SW)
##Ha: Mean(SE) != Mean(SW)
chisq.test(data_1$SE,data_1$SW)
##Since, p-value > 0.05, we accept our null hypothesis.
##Since for all regions, p-value > 0.05, the brand's share is evenly distributed
```

Answer2:

```
##Loading dataset survey
data()
View(survey)
data("survey")
print(str(survey))

##EDA
head(survey)

##Dataframe
survey_1 <- data.frame(survey)
survey_2 <- data.frame(survey_1$Smoke,survey_1$Exer)
survey_table <- table(survey_2)
survey_table

##Chi-square
##H0: Mean(smoke) = Mean(exer)
##Ha: Mean(smoke) != Mean(exer)
chisq.test(survey_table)
##since,p-value >0.05, we accept the null hypothesis.

##Correlation
cor(survey_table)
##They have strong correlation, and thus are dependent.
```

##Answer3:

```
##Loading dataset
data()
View(ChickWeight)
data("ChickWeight")
print(str(ChickWeight))
```

```
##EDA
head(ChickWeight)

##Dataframe
ChickWeight_1 <- data.frame(ChickWeight)
ChickWeight_2 <- data.frame(ChickWeight_1$Diet,ChickWeight_1$weight)
ChickWeight_2$weightcategory <-
ifelse(ChickWeight_2$ChickWeight_1.weight<median(ChickWeight_2$ChickWeight_1.weight),
"Low","High")
ChickWeight_table <-
table(ChickWeight_2$weightcategory,ChickWeight_2$ChickWeight_1.Diet)
ChickWeight_table

##Chi-square test
##H0: Mean(Diet) = Mean(Weight)
##Ha: Mean(Diet) != Mean(weight)
chisq.test(ChickWeight_table)
##Since, p-value < 0.05, we reject our null hypothesis

##Correlation
cor(ChickWeight_table)
##They are correlated, i.e. dependent.
```

##Answer4:

```
##Loading data
View(Arthritis)
data("Arthritis")
print(str(Arthritis))

##EDA
head(Arthritis)

##Dataframe
last1 <- data.frame(Arthritis)
last2 <- data.frame(last1$Treatment,last1$Improved)
last_table <- table(last2)
last_table

##Chi-Square test
##H0: Mean(Treatment) = Mean(Improved)
##H1: Mean(Treatment) != Mean(Improved)
chisq.test(last_table)
##X-squared = 13.055, p-value = 0.001463
##Since, p-value < 0.05, we reject our null hypothesis.
```