Test 1 - Solutions

Econometrics 3112.003, Spring 2019
Belk College of Business, UNCC
Ercan Karadas

[1] The joint probability distribution of T and G:

$T(row)/G(col.)$	0	5	10
5	.30	.18	.12
10	.15	.09	.06
15	.05	.03	.02

a) The marginal probability distributions of T and G:

T	5	10	15
P(T)	.6	.3	.1

G	0	5	10
P(G)	.5	.3	.2

b) Calculate $E(T)$ and $E(G)$.

$$E(T) = \sum_{t \in \{5,10,15\}} t \times P(T = t) = 5 \times 0.6 + 10 \times 0.3 + 15 \times 0.1 = 7.5$$

$$E(G) = \sum_{g \in \{0,5,10\}} g \times P(G = g) = 0 \times 0.5 + 5 \times 0.3 + 10 \times 0.2 = 3.5$$

c) Calculate the conditional probability $P(G = 10|T = 15)$.

$$P(G = 10|T = 15) = \frac{P(G = 10, T = 15)}{P(T = 15)} = \frac{0.02}{0.10} = \frac{1}{5}$$

d) Calculate $E(T|G = 5)$.
The conditional probability distributions of $T|G = 5$:

T	5	10	15
$P(T\|G = 5)$	.18/.30	.09/.30	.03/.30

Therefore,

$$\begin{aligned} E(T|G = 5) &= \sum_{t \in \{5,10,15\}} t \times P(T = t|G = 5) \\ &= 5 \times (.18/.30) + 10 \times (.09/.30) + 15 \times (.03/.30) \\ &= 7.5 \end{aligned}$$

e) Calculate $P(T = 10 \mid T + G = 15)$.

$$P(T = 10 | T + G = 15) = \frac{0.09}{0.05 + 0.09 + 0.12} = 0.346$$

[2] Suppose that X is a random variable that takes on the values 0, 1 and 2 with the following probabilities

$$P[X = 0] = 0.3, \quad P[X = 1] = 0.20, \quad P[X = 2] = 0.50$$

Suppose we take two independent draws from this distribution, denoted by X_1 and X_2, and then compute the sample average $\overline{X} = (X_1 + X_2)/2$.

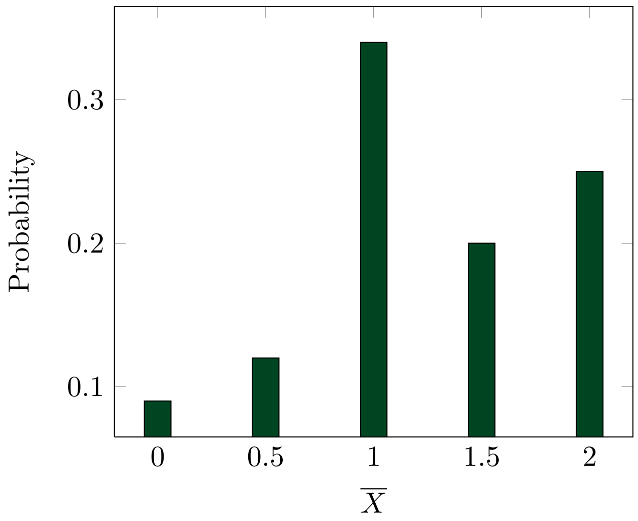
a) Since draws are independent and coming from the same distribution, they have the same expected value and variance and I will denote these by $E[X]$ and $Var[X]$, respectively.

$$\begin{aligned} E(X) &= \sum_{x \in \{0,1,2\}} x \times P(X = x) \\ &= 0 \times 0.3 + 1 \times 0.2 + 2 \times 0.5 \\ &= 1.2 \\ \text{Var}(X) &= \sum_{x \in \{0,1,2\}} (x - E(X))^2 \times P(X = x) \\ &= (0 - 1.2)^2 \times 0.3 + (1 - 1.2)^2 \times 0.2 + (2 - 1.2)^2 \times 0.5 \\ &= 0.76 \end{aligned}$$

b) The sampling distribution of $\overline{X}$:

$\overline{X}$	$P(\overline{X})$	
0	$(0.3)^2$	$= 0.09$
0.5	$2 \times (0.3)(0.2)$	$= 0.12$
1	$(0.2)^2 + 2 \times (0.3)(0.5)$	$= 0.34$
1.5	$2 \times (0.2)(0.5)$	$= 0.20$
2	$(0.5)^2$	$= 0.25$

c) Probability distribution of $\overline{X}$:



d) It doesn't look like the normal distribution and the CLT doesn't apply because the sample size is too small here ($n = 2$).

[3] The random variable Y has a mean of 30 and a variance of 16 and let $T = \frac{1}{2}Y + 3$.

a) Compute $E(T)$ and $Var(T)$.

$$E(T) = \frac{1}{2}E(Y) + 3 = 18$$
$$\text{Var}(T) = \left(\frac{1}{2}\right)^2 \text{Var}(Y) = 4$$

b) Can you compute the following probability $P(8.5 \leq T \leq 10)$? Explain why. No, we can't compute this probability because we don't know the distribution of T. Knowing $E(T)$ and $Var(T)$ doesn't mean that we know the full distribution. We have seen in class examples that two r.v.s sharing the same expected value and variance but having very different probability distributions.

Since $Y \sim N(30, 16)$, now we can say that $T \sim N(18, 4)$.

c) Compute $P(T > 21)$?

$$\begin{aligned} P(T > 21) &= P\left(Z > \frac{21 - 18}{\sqrt{4}}\right) \\ &= 1 - \Phi(1.5) \\ &= 1 - 0.9332 = 0.0668 \end{aligned}$$

d) Compute $P(14 \leq T \leq 20)$?

$$
\begin{aligned}
P(14 \le T \le 20) &= P\left(\frac{14-18}{\sqrt{4}} \le Z \le \frac{20-18}{\sqrt{4}}\right) \\
&= \Phi(1.00) - \Phi(-2.00) \\
&= 0.8413 - 0.0228 = 0.8185
\end{aligned}
$$

R questions:

```
# Problem 4
mydata <- c(-20, -15, -5, 8, 12, 9, 2, 23, 19)
#4.a
sum(x)
#4.b
mean(mydata)
#14.c
sum(mydata)/length(mydata)
#4.d.
sum(mydata[mydata>0])
#4.e.
mean(mydata[which(mydata != max(mydata))])
#4.f  Elements of mydata that are less than 8, but greater than or equal
    to 8 in absolute value, so they should be -20, -15.

# Problem 5
sumx <- mydata$x1 + mydata$x2
meanx <- sumx/2
or
meanx <- (mydata$x1 + mydata$x2)/2

# Problem 6
> vect
 foo  bar norf
  11    2   NA

# Problem 7
my_matrix <- matrix(rnorm(6), nrow=3, ncol=2)

# Problem 8
x1 <- rnorm(10)
x2 <- rnorm(10)
myDataFrame <- data.frame(a= x1, b = x1+x2)
myDataFrame

# Problem 9
> x
 [1]  0.5  1.0  1.5  2.0 25.0 30.0 35.0 40.0  4.5  5.0

# Problem 10
Y <- Orange$circumference
X <- sqrt(Orange$age)
Y + X
```