IUE – MATH 280 – Introduction to Probability and Statistics

Second Midterm Exam — December 15, 2014 — 19:00 – 20:50

Name Surname : ____________________ ID # : ____________

Q1	Q2	Q3	Q4	Q5	TOTAL
20	20	20	20	20	100

- The exam consists of **5** questions.
- Please read the questions carefully.
- Write your answers in the empty space at the end of each question. Be neat.
- Show all your work. Answers without sufficient explanation might not get full credit.
- Exchange of any material (e.g., calculators, rubbers, tables, etc.) is not allowed.
- Dictionaries and mobile phones are not allowed.
- You may use any empty space for scratch work.
- Don't forget to **indicate** your instructor and your section in the following table.

Instructor

Guvenc Arslan	ITF	Mon 17-18 & Wed 12-15	
Femin Yalcin Gulec	BA1	Mon 18-19 & Tue 11-14	
Femin Yalcin Gulec	BA2	Tue 10-11 & Thu 16-19	
Femin Yalcin Gulec	ITF	Mon 16-17 & Wed 16-19	
Femin Yalcin Gulec	LM	Mon 17-18 & Tue 16-19	

GOOD LUCK!

Question 1. Delivery trucks arrive at the distribution center of GF Company with various consumer items from company's suppliers. The mean number of trucks arriving per hour is 3.

a) What is the probability that at least 2 will arrive in the next hour?
b) What is the probability that at most 3 trucks will arrive in 20 minutes?
c) Given that a truck has just arrived, what is the probability that the next truck will arrive between 10 and 30 minutes?
d) What is the mean time (**in minutes**) between the successive arrivals of the trucks?

Solution:

a) X - number of trucks arriving at the distribution center of GF Company per hour
$X \sim$ Poisson(3)
$P(x) = P\{X = x\} = \frac{e^{-3}3^x}{x!}, \quad x = 0, 1, 2, \ldots$

$$\begin{aligned} P\{X \geq 2\} &= 1 - P\{X < 2\} = 1 - [P(0) + P(1)] = 1 - \left[\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!}\right] \\ &= 1 - 4e^{-3} = 0.8008 \end{aligned}$$

b) Y - number of trucks arriving at the distribution center of the company per 20-minute intervals
$Y \sim$ Poisson($\frac{1}{3}3$)=Poisson(1)
$P(y) = P\{Y = y\} = \frac{e^{-1}1^y}{y!} = \frac{1}{ey!}, \quad y = 0, 1, 2, \ldots$

$$\begin{aligned} P\{Y \leq 3\} &= P(0) + P(1) + P(2) + P(3) \\ &= \frac{1}{e}\left[\frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!}\right] = \frac{1}{e}\left[1 + 1 + \frac{1}{2} + \frac{1}{6}\right] = \frac{16}{6e} = 0.9810 \end{aligned}$$

c) T - the time (in hours) between successive delivery truck arrivals to the distribution center of GF Company
$T \sim$ exponential(3)
$F(t) = P\{T \leq t\} = 1 - e^{-3t}, \quad t \geq 0$

$$\begin{aligned} P\{10 \text{ minutes} < T < 30 \text{ minutes}\} &= P\left\{\frac{1}{6} \text{ hours} < T < \frac{1}{2} \text{ hours}\right\} \\ &= F\left(\frac{1}{2}\right) - F\left(\frac{1}{6}\right) \\ &= 1 - e^{-3\frac{1}{2}} - \left[1 - e^{-3\frac{1}{6}}\right] = e^{-0.5} - e^{-1.5} = 0.3834 \end{aligned}$$

d) $E(T) = \frac{1}{3}$ hours $= 20$ minutes

Question 2. Pizza FG has found that service times for orders follow a normal distribution with a mean of 40 minutes and a standard deviation of 10 minutes. A random sample of five orders was taken.

a) The probability is 0.10 that the sample mean of service times is less than how many minutes?

b) What is the probability that the sample standard deviation of service times is more than 15.4 minutes?

Solution:

X - service times for orders of Pizza FG

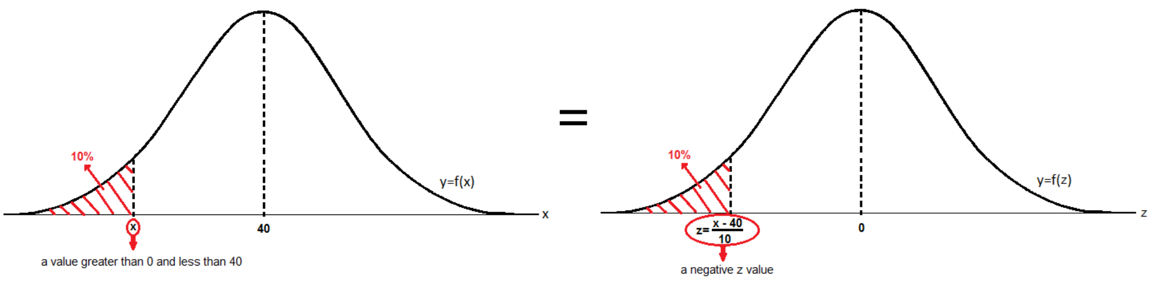
$X \sim \mathcal{N}(40, 100)$

$n = 5$

a) $P\left\{\bar{X} < x\right\} = 0.10$

$$P\left\{\bar{X} < x\right\} = P\left\{\frac{\bar{X}-40}{\frac{10}{\sqrt{5}}} < \frac{x-40}{\frac{10}{\sqrt{5}}}\right\} = P\left\{Z < \frac{x-40}{\frac{10}{\sqrt{5}}}\right\}$$

Note that



Hence

$$\begin{aligned} P\left\{\bar{X} < x\right\} &= P\left\{Z < \frac{x-40}{\frac{10}{\sqrt{5}}}\right\} = P\left\{Z > -\frac{x-40}{\frac{10}{\sqrt{5}}}\right\} = P\left\{Z > \frac{40-x}{\frac{10}{\sqrt{5}}}\right\} \\ &= 1 - P\left\{Z < \frac{40-x}{\frac{10}{\sqrt{5}}}\right\} = 1 - F\left(\frac{40-x}{\frac{10}{\sqrt{5}}}\right) = 0.10 \end{aligned}$$

$\Rightarrow F\left(\frac{40-x}{\frac{10}{\sqrt{5}}}\right) = 0.90 \approx 0.8997 \quad \Rightarrow \quad z = \frac{40-x}{\frac{10}{\sqrt{5}}} = 1.28$

$\frac{40-x}{\frac{10}{\sqrt{5}}} = 1.28 \quad \Rightarrow \quad 40 - x = 5.72 \quad \Rightarrow \quad x = 34.28 \approx 35$

b)

$$\begin{aligned} P\left\{s > 15.4\right\} &= P\left\{s^2 > (15.4)^2\right\} = P\left\{\frac{(5-1)\,s^2}{100} > \frac{(5-1)\,(15.4)^2}{100}\right\} \\ &= P\left\{\frac{4s^2}{100} > \frac{4\,(15.4)^2}{100}\right\} = P\left\{\chi_4^2 > \frac{4\,(15.4)^2}{100}\right\} = P\left\{\chi_4^2 > 9.4864\right\} \\ &\approx 0.05 \text{ (or is between 0.05 and 0.1 but very close to 0.05)} \end{aligned}$$

Question 3. It is known that the grades of a course in Statistics at IUE has a mean of 70 with a variance of 25.

a) Use Chebyshev's theorem to determine the percent of grades between 55 and 85.
b) Find an interval of grades that will include at least 75% of all the grades for this course.

Solution:
$\mu = 70$ and $\sigma^2 = 25$

a) $\mu \pm k\sigma = 70 \pm 5k = [55, 85] \Rightarrow k = 3$
According to Chebyshev's theorem <u>at least</u> $100\left[1 - \frac{1}{3^2}\right] = 88.89\%$ of all grades are between 55 and 85.

b) $100\left[1 - \frac{1}{k^2}\right] = 75\% \Rightarrow 1 - \frac{1}{k^2} = 0.75 \Rightarrow \frac{1}{k^2} = 0.25 = \frac{1}{4} \Rightarrow k^2 = 4 \Rightarrow k = 2$
The interval of grades that will include at least 75% of all grades for Statistics course is $\mu \pm k\sigma = 70 \pm 2\,(5) = [60, 80]$.

Question 4. It is important for airlines (airline companies) to follow the scheduled departure times of flights. Suppose that one airline recently sampled the records of 250 flights departing from Izmir Adnan Menderes Airport found that 3 flights were delayed for severe weather, 2 were delayed for maintenance concerns, and all the other flights were on-time.

a) Estimate the proportion of on-time departures using a 94% confidence level.
b) Estimate the proportion of on-time departures using a 99% confidence level.
c) **Interpret** and **compare** your findings in parts a) and b).

Solution:
Let $\hat{p}$ be the proportion of the flights of that airline departed from Izmir Adnan Menderes Airport **on-time** in the sample of 250 flights.
Then $n = 250$ and $\hat{p} = \frac{250-3-2}{250} = \frac{245}{250} = 0.98$.

a) $100\,(1-\alpha)\,\% = 94\% \Rightarrow 1-\alpha = 0.94 \Rightarrow \alpha = 0.06 \Rightarrow \frac{\alpha}{2} = 0.03$
RF$= z_{\frac{\alpha}{2}} = z_{0.03} = 1.88$
SE$= \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.98(0.02)}{250}} = 0.0089$
ME$=$ (RF)(SE) $= (1.88)\,(0.0089) = 0.0167$
LCL$= \hat{p} - \text{ME} = 0.98 - 0.0167 = 0.9633$
UCL$= \hat{p} + \text{ME} = 0.98 + 0.0167 = 0.9967$

b) $100\,(1-\alpha)\,\% = 99\% \Rightarrow 1-\alpha = 0.99 \Rightarrow \alpha = 0.01 \Rightarrow \frac{\alpha}{2} = 0.005$
RF$= z_{\frac{\alpha}{2}} = z_{0.005} = 2.576$
SE$=$ 0.0089 (SE didn't change because n and $\hat{p}$ didn't change.)
ME$=$ (RF)(SE) $= (2.576)\,(0.0089) = 0.0229$
LCL$= \hat{p} - \text{ME} = 0.98 - 0.0229 = 0.9571$
UCL$= \hat{p} + \text{ME} = 0.98 + 0.0229 = 1.0029$

c) For part a) we can say that a 94% confidence interval for the proportion of on-time departures is $[0.9633, 0.9967]$.
Similarly, for part b), we can say that a 99% confidence interval for the proportion of on-time departures is $[0.9571, 1]$, which is a wider interval than the previous one. (Notice that the upper confidence limit is not 1.0029 because **a proportion cannot exceed 1**.)
As a result, the confidence interval gets wider as the confidence level gets higher.

Question 5. Twelve government employees working for the ministry of revenue verifying personal income tax returns were randomly chosen to attend a one-day conference on how to increase productivity. The following table contains the average number of income tax returns verified per day for each of the twelve employees before and after attending the conference.

Before	9	11	12	13		10	12	8	5	9	15	9
After	10	11	11	14	8	10	11	9		11	14	9

a) Assuming that the table entries are normally distributed, construct a 98% confidence interval for the true mean difference between an employee's productivity before and after attending the conference.

b) Depending on your results in part a), can we make any comment on the effectiveness of this one-day conference? If yes, is it effective or not?

Solution:

Before (x_i)	After (y_i)	$d_i = y_i - x_i$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$
9	10	1	0.8	0.64
11	11	0	-0.2	0.04
12	11	-1	-1.2	1.44
13	14	1	0.8	0.64
10	10	0	-0.2	0.04
12	11	-1	-1.2	1.44
8	9	1	0.8	0.64
9	11	2	1.8	3.24
15	14	-1	-1.2	1.44
9	9	0	-0.2	0.04
TOTAL		2	0	9.6

$\bar{d} = \frac{2}{10} = 0.2$ and $s_d^2 = \frac{9.6}{9} = 1.0667 \Rightarrow s_d = 1.0328$

a) $100\,(1-\alpha)\,\% = 98\% \Rightarrow 1-\alpha = 0.98 \Rightarrow \alpha = 0.02 \Rightarrow \frac{\alpha}{2} = 0.01$

RF$= t_{n-1,\frac{\alpha}{2}} = t_{9,0.01} = 2.821$

SE$= \frac{s_d}{\sqrt{n}} = \frac{1.0328}{\sqrt{10}} = 0.3266$

ME$=$ (RF)(SE) $= (2.821)\,(0.3266) = 0.9213$

LCL$= \bar{d} - \text{ME} = 0.2 - 0.9213 = -0.7213$

UCL$= \bar{d} + \text{ME} = 0.2 + 0.9213 = 1.1213$

b) A 98% confidence interval for the true mean difference between an employee's productivity level after and before attending the conference is $[-0.7213, 1.1213]$, which means that we **cannot** make any comment on the effectiveness of this one-day conference.