Midterm 2

PSTAT 5A, Summer B, 2018

Name:

Perm #:

Section Time & TA Name (draw a circle):	A. Bernstein TW 2 pm	A. Bernstein TW 3 pm	K. Wang TW 5 pm

Instructions:

- You have 80 minutes to complete the exam.
- Read each question carefully and answer all questions.
- **Round numbers to 3 decimal places.**
- You must show your work clearly: NO WORK=NO CREDIT.
- Anyone found copying another students' work will be given an F for the course.
- You are **NOT ALLOWED** to consult any notes or textbook during this exam.
- You are **NOT ALLOWED** to consult any cellphones, smartphones, computers or electronic device of any form during this exam.
- All cellphones, smartphones and computers must be turned off.
- You may use a calculator. You cannot not use a phone as a calculator.

Questions	Points
1 (40 pts)	
2 (30 pts)	
3 (30 pts)	
TOTAL	

Good Luck!!!

Question 1

Let $\boldsymbol{Z}$ be a standard normal random variable. You sample this variable and you record its value each time. Based on the value of $\boldsymbol{Z}$, you assign a value to $\boldsymbol{W}$ as follows:

$$\boldsymbol{W} = \begin{cases} 2 & \text{if } \boldsymbol{Z} \leq -\mathbf{0.51} \\ 4 & \text{if } \boldsymbol{Z} \geq \mathbf{0.51} \\ 7 & \text{otherwise} \end{cases}$$

a) (10 points) What's the probability that $\boldsymbol{W}$ is even?

$$P(W\ is\ even) = P(W = 2) + P(W = 4)$$
$$P(W = 2) = P(\boldsymbol{Z} \leq -\mathbf{0.51}) = 0.305$$
$$P(W = 4) = P(\boldsymbol{Z} \geq \mathbf{0.51}) = 1 - P(\boldsymbol{Z} \leq \mathbf{0.51}) = 1 - 0.695 = 0.305$$

Hence: $P(W\ is\ even) = P(W = 2) + P(W = 4) = 0.305 + 0.305 = 0.610$

b) (15 pts) You sample $\boldsymbol{Z}$ ten thousand times and you record all the values. What's the probability that $\boldsymbol{W}$ turns out odd between 4000 and 5000 (inclusive) times? Find an approximate value.

This is a Binomial Experiment. Let $X \sim Bin(n, p)$ with $n = 10{,}000$ and $p = P(W\ is\ odd) = 1 - P(W\ is\ even) = 1 - 0.610 = 0.39$

Let $Y \sim N(\mu, \sigma)$ where $\mu = np = 100 \times 0.39 = 3900$ and $\sigma = \sqrt{np(1-p)} = \sqrt{100 \times 0.39 \times (1 - 0.39)} = 48.775$

Next, use the binomial approximation of X with Y:

$$P(4000 \leq X \leq 5000) = P(3999.5 < X < 5000.5) \approx P(3999.5 < Y < 5000.5)$$
$$= P(Y < 5000.5) - P(Y < 3999.5)$$
$$P(Y < 5000.5) = P\left(Z < \frac{5000.5 - 3900}{48.775}\right) = P(Z < 22.563) = 1$$
$$P(Y < 3999.5) = P\left(Z < \frac{3999.5 - 3900}{48.775}\right) = P(Z < 2.040) = 0.9793$$

Finally:

$P(4000 \leq X \leq 5000) \approx P(3999.5 < Y < 5000.5) = 1 - 0.9793 = 0.0207$

c) (10 pts) Find the approximate probability that $\boldsymbol{W}$ turns out even 6000 times.

This is a Binomial Experiment. Let $X \sim Bin(n, p)$ with $n = 10{,}000$ and $p = P(W\ is\ even) = 0.610$

Let $Y \sim N(\mu, \sigma)$ where $\mu = np = 100 \times 0.61 = 6100$ and

$\sigma = \sqrt{np(1-p)} = \sqrt{100 \times 0.61 \times (1 - 0.61)} = 48.775$

Next, use the binomial approximation of X with Y:

$P(X = 6000) = P(5999.5 < X < 6000.5) \approx P(5999.5 < Y < 6000.5)$
$= P(Y < 6000.5) - P(Y < 5999.5)$

$$P(Y < 6000.5) = P\left(Z < \frac{6000.5 - 6100}{48.775}\right) = P(Z < -2.04) = 0.0207$$
$$P(Y < 5999.5) = P\left(Z < \frac{5999.5 - 6100}{48.775}\right) = P(Z < -2.06) = 0.0197$$

Finally:

$P(X = 6000) \approx P(5999.5 < Y < 6000.5) = 0.0207 - 0.0197 = 0.001$

d) (5 pts) Check any conditions used in b) or c) for your approximation to hold.

For b):

$$np = 10000 \times 0.39 = 3900 \geq 10$$
$$n(1-p) = 10000 \times (1 - 0.39) = 6100 \geq 10$$

For c):

$$np = 10000 \times 0.61 = 6100 \geq 10$$
$$n(1-p) = 10000 \times (1 - 0.61) = 3900 \geq 10$$

Question 2

You graduate from UCSB and your first job is to manage the Lusail Iconic Stadium for the 2022 Qatar FIFA World Cup with capacity for 86,250 people. As a stadium manager, you believe that estimating the number of people attending a match is crucial for your decisions. To start working on this, you analyze data from this year's FIFA World Cup and obtain the following values:

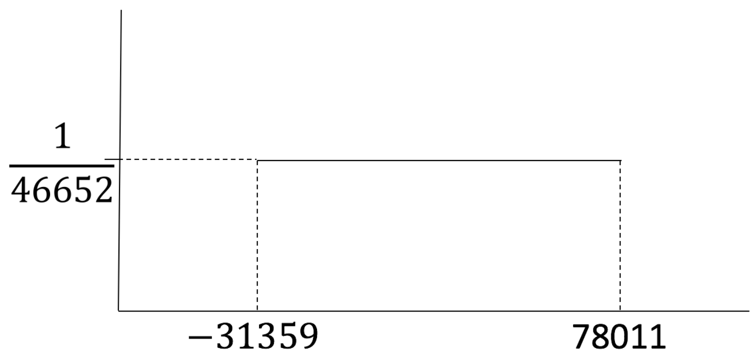
# of matches	Average	St. Dev.	Min	Max
64	45363	13000	31359	78011

At this point in the question, you realize that the table describes sample data. And its entries should be read as:

n	$\bar{x}$	s	Min	Max
64	45363	13000	31359	78011

a) (10 pts) At first you believe that attendance to any match is uniform. What's the probability that a match has an attendance between 70000 and 80000 people?

$\hat{a} = 31359$; $\hat{b} = 78011$ so use $U \sim Unif(a = 31359, b = 78011)$



$$P(70000 < U < 80000) = P(70000 < U < 78011) = \frac{78011 - 70000}{46652} \approx 0.172$$

After a few days of examining the data you realize that it actually looks more like a Normal Distribution.

b) (10 pts) The FIFA President emails you. He wants a 98% confidence interval for the mean match attendance. What would you report to him?

σ is unknown, so use : $\bar{x} \pm t\frac{s}{\sqrt{n}}$

$df = n - 1 = 64 - 1 = 63$ then $t = 2.387$

Then: $\bar{x} \pm t\frac{s}{\sqrt{n}} = 45363 \pm 2.387\ \frac{13000}{\sqrt{64}} = 45363 \pm 3878.875$
$= (41484.12, 49241.88)$

c) (10 pts) A week later, the FIFA President tells the press that "the FIFA World Cup is a crowded event with a regular match having at least 50000 people attending". Would you agree with him at a 1% significance level? The FIFA President thinks the real standard deviation is actually 12000.

$H_0: \mu \geq 50{,}000$ vs $H_A: \mu < 50{,}000$
$\sigma = 12{,}000$
$\alpha = 0.01$

$$z_{\text{test}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{45363 - 50000}{\frac{12000}{\sqrt{64}}} = -3.091$$

Since: $H_A: \mu < \mu_0 \Rightarrow$ p-value $= P(Z < z_{\text{test}}) = P(Z < -3.091) = 0.001$

Finally: p-value<α since $0.001 < 0.01$, then reject H_0

Conclusion: We reject the null hypothesis that a regular match has at least 50000 people attending. Hence, we I don't agree with the FIFA President.

Question 3

Continuing with the stadium manager job and same data from Question 2:

# of matches	Average	St. Dev.	Min	Max
64	45363	13000	31359	78011

This time you don't believe the FIFA President anymore. Forget everything he said or claimed.

a) (10 pts) Relying only on the data and your point estimates, what is the 80th percentile of people attendance for a single match?

We use the percentile formula seen in class: $x_{80} = \mu + \sigma z_{80}$

Based solely on the data, we use the point estimates:

$$\hat{\mu} = \bar{x} = 45363$$
$$\hat{\sigma} = s = 13000$$

The 80th percentile for a standard normal is: $z_{80} = 0.84$ from the Z-table

Then:

$$x_{80} = 45363 + 13000 \times 0.84 = 56283$$

b) (10 pts) You would like to have a confidence interval with a margin of error of 3000 at a 95% confidence level. According to theory, how many more observations of World Cup matches would you need to achieve that?

We need to compute the minimum sample size for a margin of error using:

$n_{MIN} = \left(\frac{ts}{E}\right)^2$ because σ is unknown.

$E = 3000$

$df = n - 1 = 63$ at a 95% confidence level, hence: $t = 1.998$

$n_{MIN} = \left(\frac{ts}{E}\right)^2 = \left(\frac{1.998 \times 13000}{3000}\right)^2 = 74.96$ which is not an integer, so we roundit up to the next integer: 75

Then, according to the theory, we would need additional $75 - 64 = 11$ observations

c) (10 pts) Can the mean match attendance be 48000 at a 5% significance level? Assume the standard deviation is actually 10900.

$H_0: \mu = 48{,}000$ vs $H_A: \mu \neq 48{,}000$
$\sigma = 10{,}900$
$\alpha = 0.05$

$$z_{\text{test}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{45363 - 48000}{\frac{10900}{\sqrt{64}}} = -1.935$$

Since: $H_A: \mu \neq \mu_0 \Rightarrow$
p-value $= P(Z < z_{\text{test}}) = P(Z < -1.935) + P(Z > 1.935)$=2×0.0314 $= 0.0628$
From the Z-table: $P(Z < 1.93) = 0.9732$ and $P(Z < 1.94) = 0.9738$. So :

$$P(Z < 1.935) \approx \frac{0.9732 + 0.9738}{2} = 0.9735$$

Hence: p-value $= 0.0265$ + (1-$0.9735) = 2 \times 0.0265 = 0.053$
Finally: p-value>α since $0.053 > 0.05$, then we can't reject H_0

Conclusion: We can't reject the null hypothesis that the mean attendance to a regular match is 48000 people