Midterm #1

answers

Instructions: The exam is divided into two parts. In Part I, your score is entirely based on your numerical answer, and not any accompanying explanation. In Part II, your grade for each question will be based primarily on your written explanation, and unsupported answers will receive no points. For both parts, you should refer to the provided probability tables as necessary. You may use a calculator, and one sheet of notes.

Part I (8 questions)

Each question is worth 5 points. Write your numerical answer in the space provided. Any work you show is irrelevant to your grade.

Problem 1 X is a Poisson random variable with mean 10. Find $P(X > 12)$.

.208

Problem 2 Professor Gonzo forgets to show up to 20% of his classes. If there are 25 scheduled meetings throughout the semester, what is the probability he teaches at least 20 classes (assume his attendance is independent across classes)?

.617

Problem 3 X is a normal random variable with mean 4 and standard deviation 1. What is $P(X \geq 5)$?

.159

Problem 4 X is a normal random variable with mean 200 and standard deviation 8. Find $P(X \leq 240)$.

This probability is very, very close to one (any answer that makes this point is fine).

Problem 5 The number of daily newspapers sold by a newstand is normally distributed with mean 1,000 and standard deviation 100. How many papers should the stand stock so that it sells out on no more than 40% of days?

1,026 newspapers

Problem 6 X is a normal random variable with mean 5 and standard deviation 1. Find $P(-3 < X < 7)$.

.977

Problem 7 You take a sample of size 49 from a population with mean $\mu = 100$ and standard deviation $\sigma = 21$. What is the probability the sample mean is greater than 100?

$\frac{1}{2}$

Problem 8 You take a sample of size 64 from a population with mean $\mu = 500$ and standard deviation $\sigma = 800$. What is the probability the sample mean is less than 600?

.841

Part II (4 questions)

Each question is worth 15 total points. Thoroughly support all of your answers.

Problem 9 The number of customers who enter a supermarket each hour is normally distributed with a mean of 600 and a standard deviation of 200. The supermarket is open 16 hours per day. What is the probability that the total number of customers who enter the supermarket in one 16-hour day is greater than 10,000?

For the total number of customers to be above 10,000, the average number of customers in each of the 16 hours must be at least $\frac{10{,}000}{16} = 625$. Think of one day as a sample of 16 hours. The sample mean then has a normal distribution with mean equal to 600 and a standard deviation equal to $\frac{200}{4} = 50$. The probability that a normal random variable with mean 600 and standard deviation 50 is larger than 625 is .309, so there is a 30.9% chance the store has more than 10,000 customers in one 16-hour day.

Some of you also pointed out that if the number of customers entering in one hour averages 600 with standard deviation 200, the number entering in 16 hours must average 9,600 with standard deviation 3,200. This is almost, but not quite correct. To sum normal random variables, we sum both the mean and the *variance*, meaning the standard deviation of the number of customers in 16 hours is 800. I gave nearly full marks for an answer along these lines.

Problem 10 A zoologist wishes to know the average length of hibernation for grizzly bears in Alaska. She tags a random sample of 49 bears and finds that the the bears in her sample hibernate for an average of 84 days. The standard deviation for hibernation length is $\sigma = 16$

a. Calculate a 95% confidence interval for the average hibernation length across the entire population of Alaska's grizzly bears.

A 95% confidence interval has the form $\overline{x} + / - 1.96 * \frac{\sigma}{\sqrt{n}}$. Her 95% confidence interval is therefore $[79.52, 88.48]$.

b. Calculate a 70% confidence interval.

A 70% confidence interval has the form $\overline{x} + / - 1.04 * \frac{\sigma}{\sqrt{n}}$. Her 70% confidence interval is therefore $[81.62, 86.38]$.

c. Calculate a 99.5% confidence interval.

A 99.5% confidence interval has the form $\overline{x} + / - 2.81 * \frac{\sigma}{\sqrt{n}}$. Her 99.5% confidence interval is therefore $[77.58, 90.42]$.

d. The zoologist, who has no training in statistics or probability, suggests that a higher confidence level *necessarily* makes for a better estimate. Is she right? Discuss any tradeoffs associated with a higher confidence level in the context of this example.

A higher confidence level is good in that it means a lower probability of μ lying outside of the confidence interval, but the tradeoff is that the confidence interval will be wider, making your estimate less precise (alternatively, to keep the same width confidence interval as confidence level increases, you must increase the sample size, which is costly in terms of time and increased probability of being mauled by a grizzly bear). Therefore, there is no one confidence level that is "optimal" under general conditions, but a 95% confidence level is often used both because of convention and because it often seems to be a happy medium balancing the costs and benefits of a higher confidence levels.

Problem 11 Three weeks into the 2012 NFL season, three NFL teams, the Arizona Cardinals, the Atlanta Falcons, and the Houston Texans, have won each of their first three football games. Every other team in the NFL has lost at least one game.

a. Suppose that the probability that the home team wins any game is 50%, regardless of the relative talent levels of the two teams playing. What is the probability of any one team winning three games in a row?

If wins are independent events, then the probability one team wins three games in a row is $\frac{1}{2}^3 = \frac{1}{8}$

b. Given your answer to a., and again assuming that each game is won by the home team with 50% probability, about how many of the NFL's 32 total teams would you expect to win their first 3 games in a typical year?

If about 1 out of 8 teams wins its first three games, on average, then there should be about 4 3-0 teams in a typical year.

c. Do you think your answer to b overestimates or underestimates the number of undefeated NFL teams after 3 weeks in a normal year? Explain why.

The answer in b is based on the assumption that the outcomes of games and randomly decided, regardless of the relative talent level of the two teams. Although this may have been the case in some games with replacement referees this season, generally there are a handful of very talented teams that win most of their games, as well as some inferior teams that lose most of their games. The result is that we'd expect more than 4 teams to be 3-0 in a typical year. For example, if the Patriots win 70% of their games in a typical year, the probability they would start 3-0 is $.7^3 = .343$. The fact that this year there are only 3 3-0 teams can be attributed to statistical noise.

Problem 12 The number of servings of lamb and beef schwarma X consumed by each customer at Sahara is a random variable with the following distribution:

value	**probability**
0	.3
1	.5
2	.2

a. Calculate the mean and standard deviation of X.

The mean is $0 * .3 + 1 * .5 + 2 * .2 = .9$. The variance is $(0 - .9)^2 * .3 + (1 - .9)^2 * .5 + (2 - .9)^2 * .2 = .49$. The standard deviation is $\sqrt{.49} = .7$.

b. Suppose Sahara has 900 customers/day, and that Sahara prepares 800 servings of schwarma each day. What is the probability that they will sell all 800 servings on any given day?

Selling 800 servings to 900 customers means that each customer must order an average of $\frac{8}{9} = .89$ servings of schwarma. Think of one day's 900 customers as a random sample of size $n = 900$. What is the probability that the sample average is above .89? By the Central Limit Theorem, $\overline{x}$ is a normal random variable with mean .9 and standard deviation $\frac{.7}{30} = .023$. Since .89 is .435 standard deviations below the mean of .9, the probability that $\overline{x}$ will be above .89 and therefore that Sahara will sell out of their delicious lamb and beef schwarma is therefore about 66.8%.