Stat 330 (Spring 2020): Homework 8 Due: April 1, 2020

Show all of your work, and *please* staple your assignment if you use more than one sheet. Write your name, the course number and the section on every sheet. Problems marked with * will be graded and one additional randomly chosen problem will be graded.

1. * The price of a particular tablet among dealers nationwide is assumed to have a Normal distribution with mean $\mu = \$600$ and variance $\sigma^2 = 36$.

   (a) What is the probability that the tablet, chosen randomly from a dealer, will cost less than \$595?

   Answer:

$$\begin{aligned} P(X < 595) &= P\left(Z < \frac{595 - 600}{6}\right) \\ &= P(Z < -0.83) = \Phi(-0.83) = 0.2033 \end{aligned}$$

   (b) What is the probability that the tablet, chosen randomly from a dealer, will cost more than \$603?

   Answer:

$$\begin{aligned} P(X > 603) &= P\left(Z > \frac{603 - 600}{6}\right) \\ &= P(Z > 0.50) = 1 - \Phi(0.50) = 1 - 0.6915 = 0.3085. \end{aligned}$$

   (c) What is the probability that the tablet, chosen randomly from a dealer, will cost between \$595 and \$603?

   Answer:

$$\begin{aligned} P(595 < X < 603) &= P\left(\frac{595 - 600}{6} < Z < \frac{603 - 600}{6}\right) \\ &= P(-0.83 < Z < 0.50) \\ &= \Phi(0.50) - \Phi(-0.83) = 0.6915 - 0.2033 = 0.4882 \end{aligned}$$

   (d) The manufacturer doesn't want dealers to markup these tablets above the 90$^{\text{th}}$ percentile of the price distribution. Approximately, what is the 90$^{\text{th}}$ percentile of the distribution of the price of these tablets?

   Answer:

$$P(Z < z) = 0.90 \implies z \approx 1.28$$

   Hence we have

$$\frac{X - 600}{6} = 1.28 \implies X = 600 + (1.28)(6) = 607.68$$

   (e) What is the probability that the **average** price of 30 tablets, chosen randomly from a dealer, will be greater than \$603?

   Answer: The average price of 30 tablets of the same make chosen randomly from a dealer has mean $\mu = \$603$ and standard deviation $\sigma/\sqrt{n} = 6/\sqrt{30}$

   Hence the probability that the average price of 30 tablets of the same make is greater than \$603 is

$$\begin{aligned} P(\bar{X} > 610) &= P\left(Z > \frac{603 - 600}{6/\sqrt{30}}\right) \\ &= P(Z > 2.74) = 1 - \Phi(2.74) = 1 - 0.9969 = 0.0031 \end{aligned}$$

2. Consider two brands of batteries for calculators, Xmax and Y-cell, manufactured **independently** by two companies. Assume that the lifetimes of each battery has a Normal distribution. The Xmax batteries has a mean of 1000 hours with a standard deviation of 100 hours, and Y-cell batteries has a mean of 1200 hours with a standard deviation of 200 hours. *For each of the following questions, you must state the random variable you are using and the distribution assumption you make.*

(a) What is the probability that an Xmax battery will last at most 800 hours? What is this probability for a Y-cell battery?

(b) Assume, two batteries are randomly selected, one Xmax and one Y-cell. What can you say about the difference D in their lifetimes? **It turns out that D is also Normally distributed**. Find the expected value and variance of the random variable D.

(c) You notice that, on average, Y-cell batteries will last longer than Xmax batteries. But what is the exact probability that the Y-cell batteries will last longer than Xmax batteries?

Answer:

Let X = the lifetime of Xmax batteries
Y = the lifetime of Y-cell batteries.

$$X \sim \mathrm{N}(1000, 100^2) \qquad Y \sim \mathrm{N}(1200, 200^2)$$

(a)

$$\begin{aligned} P[X \leq 800] &= P\left[Z \leq \frac{800-1000}{100}\right] \\ &= P[Z \leq -2] \\ &= P[Z \geq 2] = 0.0228 \qquad \text{(from Tables)} \\ P[Y \leq 800] &= P\left[Z \leq \frac{800-1200}{200}\right] \\ &= P[Z \leq -2] = 0.0228 \end{aligned}$$

(b) Take $D = Y - X$ (could take $D = X - Y$)
This is of the form aX + bY and D is distributed Normal.

$$\begin{aligned} E[D] &= E[Y] - E[X] = 1200 - 1000 = 200 \\ Var[D] &= Var[Y] + Var[X] \qquad \text{(X and Y are independent r.v.'s)} \\ &= 200^2 + 100^2 = 50,000 \end{aligned}$$

Thus $D \sim \mathrm{N}(200, 50000)$

(c) Need

$$\begin{aligned} &P[(Y-X) > 0] (\text{or} \quad P[(X-Y) < 0]) \\ =&P\left[Z > \frac{0-200}{\sqrt{50000}}\right] \\ =&P[Z > -0.8944] \\ =&1 - P[Z \leq -0.8944] = 0.8144 \end{aligned}$$

3. An iid sample of n observations is drawn from a population with mean equal to 50 and standard deviation equal to 5. Let $\bar{X}$ be the sample mean.

(a) Given $n = 25$. Approximate $P(49 < \bar{X} < 51)$.

(b) Given $n = 100$. Approximate $P(49 < \bar{X} < 51)$.

(c) Based on your answer to (a) and (b), what happens to the probability that the sample mean is between 49 and 51 as your sample size n increases?

(d) What is the smallest sample size so that the probability that $|\bar{X} - 50| > 1$ is at most 0.05?

Answer:

(a) By CLT, $\bar{X} \sim N(50, 5^2/25)$ approximately.

$$\begin{aligned} P(49 < \bar{X} < 51) &= P\left(\frac{49-50}{5/\sqrt{25}} < \frac{\bar{X}-50}{5/\sqrt{25}} < \frac{51-50}{5/\sqrt{25}}\right) \\ &= P(-1 < Z < 1) \qquad Z \sim N(0,1) \\ &= \Phi(1.00) - \Phi(-1.00) \\ &= 0.8413 - 0.1587 = 0.6826 \end{aligned}$$

(b) By CLT, $\bar{X} \sim N(50, 5^2/100)$ approximately.

$$\begin{aligned} P(49 < \bar{X} < 51) &= P\left(\frac{49-50}{5/\sqrt{100}} < \frac{\bar{X}-50}{5/\sqrt{100}} < \frac{51-50}{5/\sqrt{100}}\right) \\ &= P(-2 < Z < 2) \qquad Z \sim N(0,1) \\ &= \Phi(2.00) - \Phi(-2.00) \\ &= 0.9772 - 0.0228 = 0.9544 \end{aligned}$$

(c) As sample size increases, the probability that $\bar{X}$ is between 49 and 51 increases. (This is because as sample size increases, the distribution of your sample mean becomes more and more concentrated around the population mean of 50)

(d)

$$\begin{aligned} P(|\bar{X} - 50| > 1) &\le 0.05 \\ P(\bar{X} - 50 < -1) + P(\bar{X} - 50 > 1) &\le 0.05 \\ P(Z < -1/(5/\sqrt{n})) + P(Z > 1/(5/\sqrt{n})) &\le 0.05 \\ 2P(Z < -\sqrt{n}/5) &\le 0.05 \\ \Phi(-\sqrt{n}/5) &\le 0.025 \end{aligned}$$

which implies $-\sqrt{n}/5 \le -1.96$ and thus $n \ge 96.04$. The smallest sample size is 97.

4. The average height of professional basketball players is around 6 feet 7 inches, and the standard deviation is 3.89 inches. Assuming Normal distribution of heights within this group,

(a) What percent of professional basketball players are taller than 7 feet?

(b) If your favorite player is within the tallest 20% of all players, what can his height be?

Answer:

4.21 The height X has Normal distribution with $\mu = 79''$ and $\sigma = 3.89''$. Using Table A4,

(a) $\boldsymbol{P}(X > 84'') = \boldsymbol{P}\left(Z > \frac{84-79}{3.89}\right) = P(Z > 1.29) = 1 - \Phi(1.29) = 1 - 0.9015 =$ $\boxed{0.0985 \text{ or } 9.85\%}$

(b) Solve the equation

$$\boldsymbol{P}(X > x) = 0.20$$

for x. We have

$$\boldsymbol{P}(X > x) = \boldsymbol{P}\left(Z > \frac{x-79}{3.89}\right) = 1 - \Phi\left(\frac{x-79}{3.89}\right) = 0.20,$$

so that $\Phi\left(\frac{x-79}{3.89}\right) = 0.80$.

From Table A4, we find that $\Phi(0.84) \approx 0.80$. Therefore (unstandardizing 0.84),

$$\frac{x-79}{3.89} = 0.84 \quad \Rightarrow \quad x = \boxed{82.3 \text{ in or } 6 \text{ ft } 10.3 \text{ in}}$$

So, the height of your favorite player can be 6'10.3" or more.

5. Suppose the installation time in hours for a software on a laptop has probability density function $f(x) = \frac{4}{3}(1 - x^3)$, $0 \le x \le 1$.

   (a) Find the probability that the software takes between 0.3 and 0.5 hours to be installed on your laptop.

   (b) Let $X_1, \ldots, X_{30}$ be the installation times of the software on 30 different laptops. Assume the installation times are independent. Find the probability that the *average* installation time is between 0.3 and 0.5 hours. Cite the theorem you use.

   (c) Instead of taking a sample of 30 laptops as in the previous question, you take a sample of 60 laptops. Find the probability that the *average* installation time is between 0.3 and 0.5 hours. Cite the theorem you use.

   Answer:

   (a)
$$P(0.3 < X < 0.5) = \frac{4}{3}\int_{0.3}^{0.5}(1 - x^3)dx = \frac{4}{3}\left(x - \frac{x^4}{4}\right)\Big|_{0.3}^{0.5} = 0.2485$$

   (b) First $E(X) = \int_0^1 4x/3 \times (1 - x^3)dx = 0.4$. $Var(X) = E(X^2) - E(X)^2 = 2/9 - (2/5)^2 = 14/225$. By the central limit theorem,

$$\bar{X}_{30} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \equiv N\left(0.4, \frac{14/225}{30}\right) \equiv N(0.4, 7/3375)$$

$$\begin{aligned}
P(0.3 < \bar{X}_{30} < 0.5) &= P\left(\frac{0.3 - 0.4}{\sqrt{7/3375}} < \frac{\bar{X}_{30} - 0.4}{\sqrt{7/3375}} < \frac{0.5 - 0.4}{\sqrt{7/3375}}\right) \\
&= P(-2.20 < Z < 2.20) \\
&= P(Z < 2.20) - P(Z < -2.20) \\
&= 0.9861 - 0.0139 \\
&= 0.9722
\end{aligned}$$

   (c) $E(X) = 0.4$. $Var(X) = 14/225$. By the central limit theorem,

$$\bar{X}_{60} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \equiv N\left(0.4, \frac{14/225}{60}\right) \equiv N(0.4, 7/6750)$$

$$\begin{aligned}
P(0.3 < \bar{X}_{30} < 0.5) &= P\left(\frac{0.3 - 0.4}{\sqrt{7/6750}} < \frac{\bar{X}_{30} - 0.4}{\sqrt{7/6750}} < \frac{0.5 - 0.4}{\sqrt{7/6750}}\right) \\
&= P(< Z <) \\
&= P(Z < 3.11) - P(Z < -3.11) \\
&= 0.9991 - 0.0009 \\
&= 0.9982
\end{aligned}$$

6. * Installation of some software package requires downloading 100 files. On average, it takes 15 seconds to download one file, with a variance of 25 sec^2. What is the (approximate) probability that the software is installed in less than than 23 minutes? (Use the Central Limit Theorem)

   Answer: Let S = sum of the 100 file download times. Then, from the Central Limit Theorem, we have $S \approx N(15 \cdot 10, 25 \cdot 10) \equiv N(1500, 2500)$
   We want
$$P(S < 1380) = P\left(Z \le \frac{1380 - 1500}{\sqrt{2500}}\right) = \Phi(-2.4) = 0.0082.$$
   There's a 0.82% chance the full download would take less than 23 minutes.