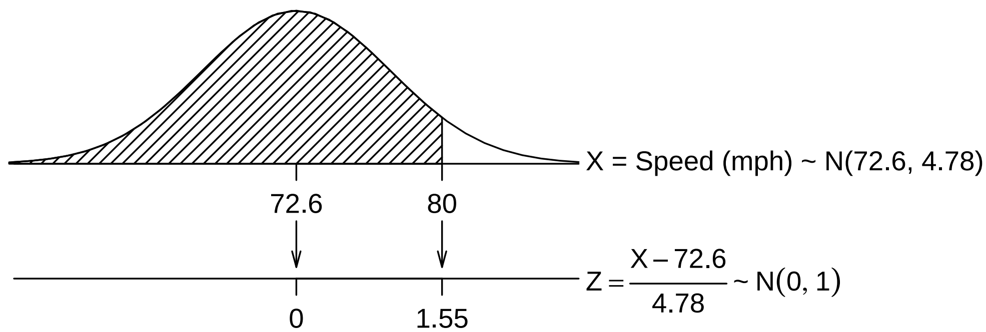
STAT22000 Summer 2020 Homework 7 Solutions

Problems to Turn In: due **midnight of Monday, July 13**, on Canvas.

1. Exercise 3.12 on p. 160 – [*8 points in total*]

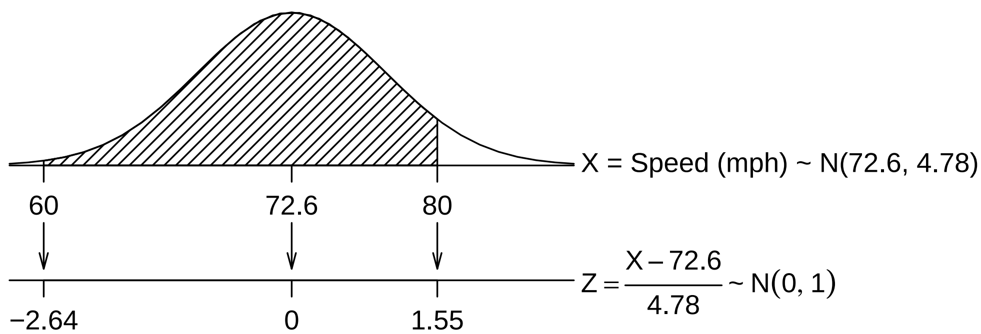
   (a) [*2pts*] Let X be the speeds of the cars traveling on this stretch of the I-5. The problem says $X \sim N(\mu = 72.6, \sigma = 4.78)$ and asks for $P(X < 80)$.



   This can be found in R using either of the following commands

```
> pnorm((80-72.6)/4.78)
[1] 0.939203
> pnorm(80, m = 72.6, s = 4.78)
[1] 0.939203
```

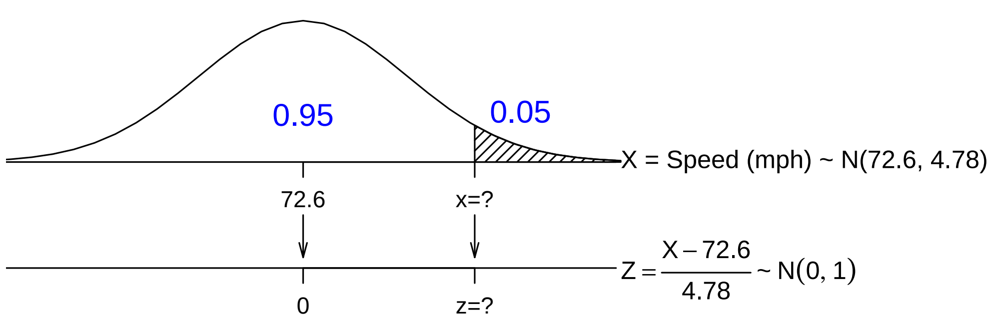
   (b) [*2pts*] The problem asks for $P(60 < X < 80) = P(X < 80) - P(X < 60)$.



   This can be found in R using either of the following commands

```
> pnorm((80-72.6)/4.78) - pnorm((60-72.6)/4.78)
[1] 0.9350083
> pnorm(80, m = 72.6, s = 4.78) - pnorm(60, m = 72.6, s = 4.78)
[1] 0.9350083
```

   (c) [*2pts*] The fastest 5% of cars travel is the 95 th percentile. The problem asks for the speed x such that $P(X > x) = 0.05$, i.e., $P(X < x) = 0.95$.

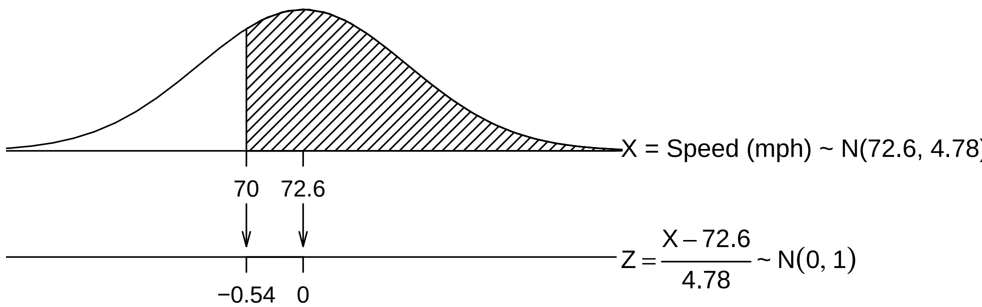


This can be found in R as follows

```
qnorm(0.95, m = 72.6, s = 4.78)
[1] 0.9350083
```

(d) [*2pts*] The problem asks for $P(X > 70)$, which can be found in R using any of the following commands to be 0.7068.

```
> pnorm((70-72.6)/4.78, lower.tail=F)
[1] 0.7067562
> 1-pnorm((70-72.6)/4.78)
[1] 0.7067562
> pnorm(70, m = 72.6, s = 4.78, lower.tail=F)
[1] 0.7067562
> 1-pnorm(70, m = 72.6, s = 4.78)
[1] 0.7067562
```



2. A student takes a multiple-choice quiz with 5 questions, each with four possible selections for the answer. A passing grade is 60% or better (i.e., answering at least 3 of 5 questions correctly). Suppose that the student was unable to find time to study for the exam and just guesses at each question. Find the probability that the student ....

   (a) gets exactly 3 questions correct.

   (b) passes the exam.

   (c) How many questions would you expect the student to get correct?

   (d) Obtain the standard deviation of the number of questions that the student gets correct.

Answer: [*6 points in total*] The number of correct guesses $X \sim Bin(n = 5, p = 1/4)$

(a) [*2pts*] $P(\text{exactly 3 questions correct}) = P(X = 3) = \binom{5}{3}(1/4)^3(3/4)^2$. As $\binom{5}{3} = \frac{5!}{3!\,(5-3)!} = \frac{5\cdot4\cdot3\cdot2\cdot1}{(3\cdot2\cdot1)(2\cdot1)} = \frac{5\cdot4}{2\cdot1} = 10$, we have $P(X = 3) = 10(1/4)^3(3/4)^2 = 45/512 \approx 0.08789$

(b) [*2pts*]

$$\begin{aligned}
P(\text{passes the exam}) &= P(X \geq 3) \\
&= P(X = 3) + P(X = 4) + P(X = 5) \\
&= \binom{5}{3}(1/4)^3(3/4)^2 + \binom{5}{4}(1/4)^4(3/4) + \binom{5}{5}(1/4)^5 \\
&= 10(1/4)^3(3/4)^2 + 5(1/4)^4(3/4) + (1/4)^5 = (90 + 15 + 1)/1024 \\
&\approx 0.08789 + 0.014648 + 0.00097656 \\
&\approx 0.1035156
\end{aligned}$$

(c) [*1pt*] The expected value is $np = 5(1/4) = 1.25$.

(d) [*1pt*] The standard deviation is $\sqrt{np(1-p)} = \sqrt{5(1/4)(3/4)} \approx 0.968$.

---

3. A fair 6-face die is going to be rolled some number of times.

   (a) Is it more likely that the ace (one spot) comes up 20% or more of the time in 60 rolls or 600 rolls? Explain. (Note that the probability that a fair 6-face die shows an ace in one roll is $1/6 = 16\frac{2}{3}\%$.)

   (b) Is it more likely to get 8 to 12 aces in 60 rolls or 98 to 102 aces in 600 rolls? Explain.

   Hint: Think about Law of Large Number (LLN). Getting an ace when rolling a fair 6-face die is like getting heads when tossing an unfair coin with only $1/6$ *probability to land heads.*

Answer: [*4 points in total, 2pts each*]

(a) 60 rolls. According to the LLN, the percentage of aces will approach $1/6 = 16\frac{2}{3}\%$ as the number of rolls increases. The percentage of aces is less likely to deviate from $16\frac{2}{3}\%$ in 600 rolls than in 60 rolls. Since you want more than 20%, you will want smaller number of rolls because this allows you more chance to deviate from $16\frac{2}{3}\%$.

(b) It's more likely to get 8 to 12 aces in 60 rolls. From the LLN, the more rolls are made, the more the number of aces deviated from 1/6 of the number of rolls ($H(n) - n/6$), and hence is more likely to be within 2 from 10 in 60 rolls than within 2 from 100 in 600 rolls.

---

4. Here we will use what we learned about the Binomial distribution to check our answer for Problem 3(b).

   (a) Find out the expected value and the standard deviation of the number of aces obtained in 60 rolls of a fair 6-face die. Ditto for 600 rolls. Find out the z-scores for the count 8 and 12 aces in 60 rolls and 98 and 102 aces in 600 rolls. Use the z-scores to explain which one is more likely.

   (b) Use the normal approximation to find the probability of getting 98 to 102 aces in 600 rolls of a fair 6-face die. Be sure to check the condition required for using this approximation. Do not use continuity correction.

   (c) Repeat the previous part but using the normal approximation with continuity correction.

   (d) Find the exact probability in (b) in R using the command below.

   ```
   pbinom(102, size=600, p=1/6)-pbinom(97, size=600, p=1/6)
   ```

   Compare the exact probability with the approximate probabilities in the previous two parts.

   Remark. The probability of getting 8 to 12 aces in 60 rolls can be calculated using the normal approximation with continuity correction or using the R command

   ```
   > sum(dbinom(8:12, size=60, p=1/6))
   [1] 0.6138631
   ```

   to be around 0.61 (not required to submit). Along with the calculation in part (b-e), you should be convinced that the probability is indeed higher with 60 rolls than with 600 rolls for the observed number of aces obtained to stay within 2 from the expected number of aces obtained.

Answer: [*7 points in total*]

(a) [2pts] The number of aces X obtained in n rolls has a binomial distribution with $\text{Bin}(n,\, p = 1/6)$. With 60 rolls, the expected value and the SD are

$$\mu = np = 60(1/6) = 10$$
$$\sigma = \sqrt{np(1-p)} = \sqrt{60(1/6)(1-1/6)} = \sqrt{50/6} \approx 2.89.$$

With 600 rolls, expected value and the SD are

$$\mu = np = 600(1/6) = 100$$
$$\sigma = \sqrt{np(1-p)} = \sqrt{600(1/6)(1-1/6)} = \sqrt{500/6} \approx 9.13.$$

The observed number of ace is less likely to deviation over 2 from the expected value with 60 rolls than with 600 rolls since the SD is smaller for 60 rolls.

(b) [2pts] We can use normal approximation to find the probability in (b) because $np = 600 \times (1/6) = 100$ and $n(1-p) = 600 \times (1-1/6) = 500$ are both ≥ 10,

$Y \sim Bin(n = 600, p = 1/6) \approx N(\mu = np = 600(1/6) = 100,\, \sigma = \sqrt{np(1-p)} = \sqrt{600(1/6)(5/6)} = \sqrt{500/6} \approx$

We can find $P(98 \leq Y \leq 102) = P(Y \leq 102) - P(Y < 98)$ using normal approximation in R as follows.

```
> pnorm(102, m = 100, s = sqrt(500/6))-pnorm(98, m = 100, s = sqrt(500/6))
[1] 0.1734193
```

(c) [2pts] With the continuity correction, the endpoints 98 and 102 becomes 97.5 and 102.5 because $P(98 \leq Y \leq 102) = P(97.5 \leq Y \leq 102.5)$ as Y is integer-valued. We can find $P(97.5 \leq Y \leq 102.5) = P(Y < 102.5) - P(Y < 97.5)$ using normal approximation in R as follows.

```
> pnorm(102.5, m = 100, s = sqrt(500/6))-pnorm(97.5, m = 100, s = sqrt(500/6))
[1] 0.2158088
```

(d) [1pt] The exact binomial probability is about 0.2157426, closer to the answer in part (d). With the continuity correction, the normal approximation gives a better approximation than without the correction.

```
> pbinom(102, size=600, p=1/6)-pbinom(97, size=600, p=1/6)
[1] 0.2157426
```
