Name:______________________________ **Date:**_______________

PSTAT 5A: Discussion Worksheet 05

Spring 2023, with *Ethan P. Marzban*

1. The weights of adult male rats is found to vary normally with mean 4.2 oz and standard deviation 0.67 oz.

(a) Define the random variable of interest, and call it X.

> **Solution:** Let X denote the height of a randomly selected adult male rat. Then, from the problem statement, $X \sim \mathcal{N}(4.2,\ 0.67)$.

(b) What proportion of adult male rats have weights under 4 oz?

> **Solution:** We seek $\mathbb{P}(X < 4)$. Since X is not *standard* normally distributed, we need to first standardize. The $z-$score associated with $x = 4$ is
>
> $$z = \frac{4 - 4.2}{0.67} \approx -0.30$$
>
> From a $z-$table, this corresponds to a probability of $\boxed{0.3821 = 38.21\%}$.

(c) What proportion of adult male rats have weights between 3.8 oz and 4.3 oz?

> **Solution:** We seek $\mathbb{P}(3.8 \leq X \leq 4.3)$. Since our table only provides left-tail areas, we first write this as
>
> $$\mathbb{P}(3.8 \leq X \leq 4.3) = \mathbb{P}(X \leq 4.3) - \mathbb{P}(X \leq 3.8)$$
>
> The associated $z-$scores are
>
> $$z_1 = \frac{4.3 - 4.2}{0.67} \approx 0.15$$
> $$z_2 = \frac{3.8 - 4.2}{0.67} \approx -0.60$$
>
> The associated probabilities (as obtained from a standard normal table) are 0.5596 and 0.2743, meaning our final answer is
>
> $$0.5596 - 0.2743 = \boxed{0.2853 = 28.53\%}$$

(d) Suppose a sample of $n = 120$ rats is taken (with replacement), and the number of rats between 3.8 oz and 4.3 oz in weight is recorded. What is the probability that this sample contains exactly 35 rats with weights between 3.8 oz and 4.3 oz? **Hint:** You will need to define another random variable and identify its distribution; you will also need your result from part (c) above.

Solution: Let Y denote the number of rats in the sample that have weight between 3.8 oz and 4.3 oz. This is actually a **binomially** distributed random variable, as can be seen by checking the 4 Binomial Criteria:

1) Independent trials? Yes, since the rats are selected *with* replacement.

2) Fixed number of trials? Yes, $n = 120$ rats.

3) Well-defined notion of "success"? Yes; "success" = "rat has weight between 3.8 and 4.3 oz"

4) Fixed probability of success? Yes; this is the quantity we found in part (c) above, so $p = 0.2853$.

Therefore, $Y \sim \mathrm{Bin}(120,\ 0.2853)$ and

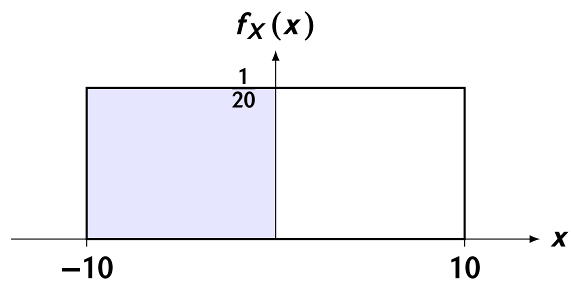
$$\mathbb{P}(Y = 35) = \binom{120}{35} \cdot (0.2853)^{35} \cdot (1 - 0.2853)^{120-35} \approx 0.07897 = 7.897\%$$

2. A particular random number generator picks a number at random from the set of real numbers between -10 and 10. The number that the generator selects can be viewed as following a uniform distribution.

(a) What is the probability that the number generated is negative?

Solution: Let X denote the number that is selected; then $X \sim \mathrm{Unif}(-10,\ 10)$. We seek $\mathbb{P}(X \leq -10)$, which we sketch as follows:

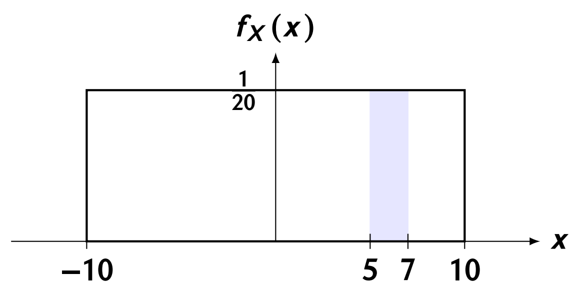
Name:______________________________ **Date:**_______________



Either by inspection, or using the formula for the area of a rectangle, we see that the area (and, consequently, the desired probability) is $0.5 = 50\%$.

(b) What is the probability that the number generated is between 5 and 7, inclusive on both ends?

Solution: We now seek $\mathbb{P}(5 \le X \le 7)$. Again, we sketch a picture:



$$\mathbb{P}(5 \le X \le 7) = \text{Area} = (7-5) \times \frac{1}{20} = \frac{2}{20} = \frac{1}{10} = 10\%$$

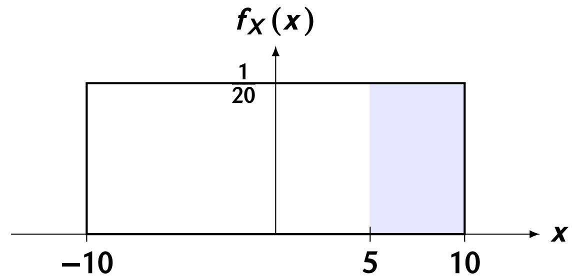
(c) What is the probability that the number generated is between 5 and 7, exclusive on both ends (i.e. *not* including 5 or 7 themselves)?

Solution: Recall that for continuous distributions, the probability that the associated random variable attains any specific value is zero. What this means is that $\mathbb{P}(5 < X < 7) = \mathbb{P}(5 \le X \le 7)$; i.e. the same result we found in part (b). 10%

(d) What is the probability that the number generated is between 5 and 12?

Solution: We now seek $\mathbb{P}(5 \le X \le 12)$. Since 12 is outside the state space of X, our picture looks like:

Name:________________________________ **Date:**________________



$$\mathbb{P}(5 \leq X \leq 12) = \text{Area} = (10 - 5) \times \frac{5}{20}\ \frac{1}{4} = 25\%$$

3. The US Census Bureau has determined that 80.7% of US Citizens live in urban areas. A representative sample of 200 US Citizens is taken, and the proportion of these that live in urban areas is recorded.

(a) Define the random variable of interest, and use the notational conventions introduced in Lecture 10.

Solution: Let $\widehat{P}$ denote the proportion of US citizens in a sample of size $n = 200$ that live in urban areas.

(b) Identify the distribution of the random variable you defined in part (a). Be sure to include the parameters of this distribution, and be sure to check any conditions that might be necessary to check!

Solution: First note that $p = 0.807$ is actually the *population* proportion of US Citizens that live in Urban Areas. Next, we check the success-failure conditions:

1) $np = (200)(0.807) = 161.4 \geq 10$ ✓

2) $n(1 - p) = (200)(1 - 0.807) = 38.6 \geq 10$ ✓

Since both conditions are satisfied, we can invoke the **Central Limit Theorem for Proportions** to say

$$\widehat{P} \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

$$\sim \mathcal{N}\left(0.807, \sqrt{\frac{0.807 \cdot (1 - 0.807)}{200}}\right) \sim \mathcal{N}(0.807,\ 0.0279)$$

(c) What is the probability that the proportion of citizens in the sample that live in urban areas lies within 2.5% of the true proportion of 80.7%?

Solution: We seek $\mathbb{P}(0.782 \leq \widehat{P} \leq 0.832)$. We find this using much the same procedure we used in Problem 1 of this worksheet: standardize, and then consult a table. First we write

$$\mathbb{P}(0.782 \leq \widehat{P} \leq 0.832) = \mathbb{P}(\widehat{P} \leq 0.832) - \mathbb{P}(\widehat{P} \leq 0.782)$$

Next, we find the associated z−scores:

$$z_1 = \frac{0.832 - 0.807}{0.0279} \approx 0.90$$
$$z_2 = \frac{0.782 - 0.807}{0.0279} \approx -0.90$$

Hence, upon consulting a table, we see that the desired probability is

$$0.8159 - 0.1841 = 0.6318 = 63.18\%$$