QUIZZ QUESTION

After being rejected for employment, Kim Kelly learns that the Bellevue Seed Company has hired only 2 women among the last 20 new employees. She also learns that the pool of applicants is very large, with an approximately equal number of qualified men and women. What is the probability of getting 2 or fewer women hired among 20, if there is no discrimination?

a) 0.0002.

b) 0.0014.

c) 0.0005.

d) 0.0173.

e) None of the above.

\[ X = \# \text{ of women hired for the last } 20 \text{ jobs} \]

\[ X \sim Bin \left( 20, \frac{1}{2} \right) \]

\[ \Pr(X \leq 2) = \Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2) \]

1) Equal \# of men and women in the pool

2) No discrimination.
\[ P_1(X = 0) = \frac{20!}{20!} \left( \frac{1}{2} \right)^0 \left( \frac{1}{2} \right)^{20} = \left( \frac{1}{2} \right)^{20} \]

\[ P_1(X = 1) = \frac{20!}{19!} \left( \frac{1}{2} \right)^1 \left( \frac{1}{2} \right)^{19} = 20 \left( \frac{1}{2} \right)^{20} \]

\[ P_1(X = 2) = \frac{20!}{18!} \left( \frac{1}{2} \right)^2 \left( \frac{1}{2} \right)^{18} = 190 \left( \frac{1}{2} \right)^{20} \]

\[ \frac{20 \cdot 19 \cdot 18!}{18! \cdot 2} \]

\[ P_1(X \leq 2) \approx 0.0002 \]
QUICK QUESTION

- Currently, an average of 7 residents of the village of Westport (population 760) die each year. What is the probability that, on a given year, there is exactly one death on a given day.
  a) 0.013.
  b) 0.009.
  c) 0.002.
  d) 0.017.
  e) None of the above

\[ X = \# \text{ of death in one day} \]
\[ X \sim \text{Poisson} \left( \frac{7}{365} \right) \]
\[ \mu = 7 \frac{\text{death}}{\text{year}} \times \mu = \frac{7}{365} \frac{\text{death}}{\text{day}} \]
\[ \Pr(X = 1) = \left( \frac{7}{365} \right)^1 e^{-7/365} \]
\[ = 0.019 \]
\[ = 1.019 \times 10^{-2} \]
QUIZZ QUESTION

- What is the expectation of the random variable whose density is depicted in the figure?
  
  a) 4  
  b) 2  
  c) 1  
  d) 3  
  e) This is not a well defined pdf.

Total area under the function

\[
= \frac{4 \cdot \frac{1}{2}}{2} = \frac{2}{2} = 1
\]
QUIZZ QUESTION

For a random variable $X$ with the density depicted in the figure, what is the probability that $X$ will be less than 1?

a) $1/2$
b) $1/4$
c) $1/8$
d) $1/6$
e) None of the above.

\[ P(X < 1) = \frac{1 \times \frac{1}{4}}{2} = \frac{1}{8} \]
NORMAL (GAUSSIAN) RANDOM VARIABLES

- The graph of the normal density is symmetric and bell-shaped. However, this is not the only density with these features.
\[ P(z \leq -2.83) = 0.0023 \]

\[ P(z \leq -1.26) = 0.1038 \]
EXAMPLE

- Let $X$ be a random variable that represents the systolic blood pressure. For the population of 18 to 74-year-old males in the US, this variable is approximately normally distributed with mean of 129 mm Hg and a standard deviation of 19.8 mm Hg.
  - Which proportion of men in the population have systolic blood pressures greater than 150 mm Hg.
  - Which pressure value represents the 5th percentile of the distribution?

$$Z \sim N(129, 19.8)$$

\[ \Pr(Z > 150) \]

\[ \Pr(Z > 150) = \Pr\left(\frac{Z - 129}{19.8} > \frac{150 - 129}{19.8}\right) \]

\[ X \sim N(0,1) \]
This just means that if
\[ X \sim N(10.7) \quad \Rightarrow \quad Z \sim N(129, 19.8) \]

\[ P(Z > 150) = P(X > \frac{150 - 129}{19.8}) \]

\[ = P(X > 1.06) \]

\[ = 0.1446 \]

\[ = 1 - P(X \leq 1.06) \]

\[ A = B \]

because the distribution is symmetric.
\[ P_X (\frac{\bar{X} - 129}{19.8} \leq \frac{9 - 129}{19.8}) = 0.05 \]

\[ P_X (X \leq \frac{9 - 129}{19.8}) = 0.05 \]

\[ \frac{9 - 129}{19.8} = -1.645 \]

\[ q = 129 - 1.645 \times 19.8 = 96.43 \]
QUESTION

- In the same example with the systolic pressure, which follows a normal distribution with mean of 129 mm Hg and a standard deviation of 19.8 mm Hg, what is the proportion of 18 to 74-year-old males in the US with blood pressure between 140 and 160 mm Hg?

  a) 0.231  
  b) 0.356  
  c) 0.174  
  d) 0.292  
  e) None of the above.

\[
Pr\left(140 \leq \frac{Z}{19.8} \leq 160 \right) = Pr\left(\frac{Z}{19.8} \leq 160\right) - Pr\left(\frac{Z}{19.8} \geq 140\right)
\]

\[
Pr\left(Z \leq 160\right) = Pr\left(X \leq \frac{160 - 129}{19.8}\right) = Pr\left(X \leq 1.56\right)
\]

\[
0.9406
\]

\[
Pr\left(Z \leq 140\right) = Pr\left(X \leq \frac{140 - 129}{19.8}\right) = Pr\left(X \leq 0.55\right)
\]

\[
0.7088
\]
EXAMPLE

○ In the same problem, what is the probability that 12 man randomly selected will have a mean weight greater than 175lb?

a) 0.360
b) 0.628
c) 0.416
d) 0.577
e) None of the above

\[ \bar{z} = \frac{\sum z_i}{n} \]

\[ \bar{z} \sim N(172, \frac{29}{127}) \]

\[ P(\bar{z} > 175) = P(x > \frac{175 - 172}{\sqrt{29/127}}) \]

\[ = P(x > \sqrt{29/127}(175 - 172)) \]

\[ = 0.360 \]
SAMPLING DISTRIBUTION OF THE MEAN

- We are interested in estimating the expectation $\mu$ of a random variable. For example, we could use one of the measures of center. In particular, we could use the average

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- Note that if you collect different samples, the mean will change slightly. Therefore the mean is a random variable. What can we say about the distribution of the mean?
SAMPLING DISTRIBUTION OF THE MEAN

- If the distribution of the underlying random variables has mean $\mu$ and standard deviation $\sigma$ then
  - The mean of the sampling distribution is $\mu$.
  - The standard deviation of the distribution of sample means is $\sigma/\sqrt{n}$.
  - If $n$ is large, the shape of the distribution of sample means is approximately normal. This result is known as the central limit theorem.
  - If the $X_i$s are Gaussian, then the central limit theorem is true for any $n$. 
EXPONENTIAL DISTRIBUTION

- A random variable $X$ follows an exponential distribution, $X \sim \text{Exp}(\mu)$, if its density is
  $$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}} \quad x \geq 0$$

- Widely used to model waiting times in line, lifetime of electronic components.
- The cdf is in this case
  $$F(x) = P(X \leq x) = 1 - e^{-\frac{x}{\mu}}$$
- The mean equals $\mu$.

(mean lifetime)
(1000 hours for a lightbulb)

\[\text{\textbullet} \quad \frac{1}{\mu} \quad f(x) \quad \text{\textbullet} \quad x\]
EXPONENTIAL DISTRIBUTION
QUIZZ QUESTION

- You are installing two lights outside your door, which requires you to go up a ladder. Each light has a lifetime that is exponentially distributed (independently of the other) with mean of 4 months. Assuming that you are a bit acrophobic (afraid of heights), and therefore you will replace the lights only when both have burned out, what is the probability that you will have to go up the ladder again within the next 4 months?

\[ \text{a) } (1 - e^{-1})^2 \]
\[ \text{b) } e^{-2} \]
\[ \text{c) } 2e^{-1} \]
\[ \text{d) } e^{-1}(1-e^{-1}) \]

\[ T_1 = \text{time before light one burns out} \]
\[ T_2 = \text{time before light two burns out} \]

\[ \text{Pr}(\text{Going up the ladder in the next 4 months}) \]
\[ = \text{Pr}(\text{both lights burn out in the next 4 months}) \]
\[ = \text{Pr}(T_1 \leq 4 \cap T_2 \leq 4) \]
\[ = \text{Pr}(T_1 \leq 4) \text{ Pr}(T_2 \leq 4) \quad \text{because of independence.} \]
I know $T_1 \sim \text{Exp}(4)$, $T_2 \sim \text{Exp}(4)$

$\Pr(T_1 \leq 4) = 1 - e^{-\frac{4}{4}} = 1 - e^{-1}$

$\Pr(T_2 \leq 4) = 1 - e^{-\frac{4}{2}} = 1 - e^{-1}$

$\Pr(\text{I will have to go up the ladder within the next 4 months})$

$= (1 - e^{-1}) (1 - e^{-1}) = (1 - e^{-1})^2$
**EXAMPLE**

- The ski gondola in Vail, Colorado, carries skiers to the top of a mountain. It bears a plaque stating that the maximum capacity is 12 people or 2100 pounds. That capacity will be exceeded if 12 people have weights with a mean great than 2100/12 = 175 pounds. Because men tend to weigh more than women, a "worst case" scenario involves 12 passengers who are all men. Men have weights that are normally distributed with mean of 172 lb and standard deviation of 29 lb. What is the probability that a single man randomly selected has a weight greater than 175 lb?
  - a) 0.628
  - b) 0.436
  - c) 0.512
  - d) 0.459
  - e) None of the above.

\[ Z = \text{weight of one man} \]

\[ Z \sim N(172, 29) \]

\[ P(Z > 175) = P(X > \frac{175 - 172}{29}) \]

\[ = 1 - 0.5398 = 0.459 \]
EXAMPLE

If you had to design a new ski gondola that carries 12 people, what would be the weight you would use for the maximum carrying size?

Say that we want the gondola to work 95% of the time. Call the maximum weight $g$

$P(X \leq g) = 0.95$

$P(X \leq \frac{g - 172}{29/\sqrt{12}}) = 0.95$

$\Rightarrow N(0,1) \Rightarrow 1.645 = \frac{g - 172}{29/\sqrt{12}}$