HSCI 50 LAB 3: Binomial and Normal Distribution ANSWER KEY

INSTRUCTIONS

The data set that we will be using for the entire course is resampled data from the MIT COVID-19 Beliefs, Behaviors & Norms Survey (https://covidsurvey.mit.edu/api.html). This is a multi-country, online survey that examined different COVID-19 perceptions across time, from July 6, 2020 to March 28, 2021. We will be using data from the Philippines aged 20 to 60.

In this lab, you will practice calculating probabilities given a binomial distribution and a normal distribution.

You have many options to submit this worksheet. Either you work on this by hand and scan/take a clear photo of your submission and save as PDF, or type your responses in a Word processor or PowerPoint presentation. You do not have to copy the questions again, but please number them accordingly.

Part A. Binomial Distribution.

We will revisit the table from Lab 2 and recode the vaccine acceptance categories into a new variable, vaccine hesitancy. Those who do not accept the vaccine or are unsure about it will be considered vaccine hesitant, otherwise they are not vaccine hesitant. The table below summarizes that information.

Table 1: COVID-19 vaccine hesitancy in the Philippines, March 14-28, 2021

Vaccine hesitant	Count	Percent
No	612	53.4
Yes	533	46.6

1. Why was dichotomizing the vaccine acceptance outcome necessary to calculate probabilities based on the binomial distribution?

Because one of the assumptions required for binomial distribution is that the outcome is binary.

2. One of the other assumptions of the binomial distribution is that the outcomes are independent from one another. This assumption cannot be tested empirically (meaning, with whatever data you have). In the simplest of explanations, what does this mean in the context of this data?

The responses of individuals to the vaccine hesitancy question is not influenced in any way by any other individual responding.

3. Another assumption of the binomial distribution is that the probabilities of the outcomes of the trials remain the same across trials. In the context of this data, what does that mean exactly?

We assume that for the probability that a person is vaccine hesitant is the same across individuals randomly selected, and that the probability is 0.466.

For questions 4-8, calculate this question by hand (you may use a calculator), and show your solutions. Approximate your answers to the nearest thousandth. Whenever appropriate, you may use the University of Iowa Binomial Distribution Applet (https://homepage.divms.uiowa.edu/~mbognar/applets/bin.html). Assume that three persons are randomly selected from this sample.

4. What is the probability that only the first person selected will be vaccine hesitant?

$$P(\text{only 1st person hesitant}) = P(\text{1st person hesitant}) \times P(\text{2nd person not hesitant}) \times P(\text{3rd person not hesitant})$$
$$= 0.466 \times 0.534 \times 0.534$$
$$= 0.133$$

5. What is the probability that only one person selected will be vaccine hesitant?

We have to consider three independent events: that the first person selected is the only one hesitant, that the second person selected is the only one hesitant, and that the third person selected is the only hesitant. We have already caclulated the probability of the first event from Question 4, and this is also the same probability for the other two events because we assume that the probability of the outcome is the same for all individuals.

Using the addition rule:

$$P(\text{only one person hesitant}) = P(\text{only 1st person hesitant}) + P(\text{only 2nd person hesitant}) + P(\text{only 3rd person hesitant})$$
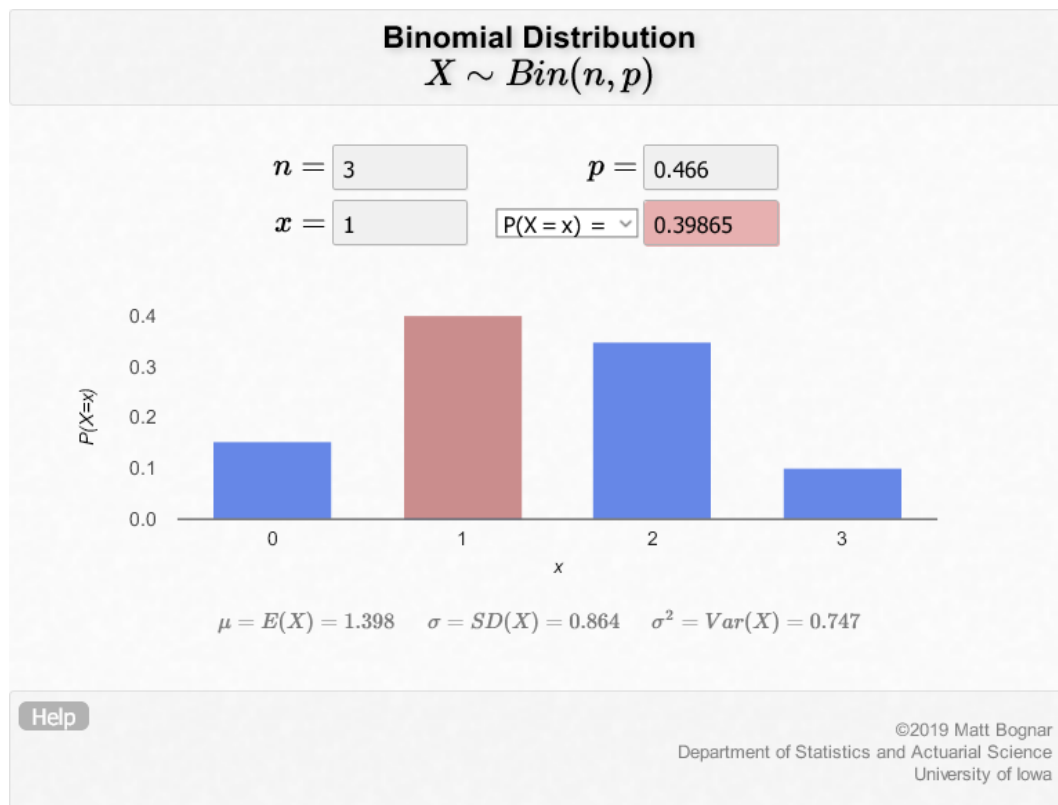$$= 0.133 + 0.133 + 0.133$$
$$= 0.399$$

You may also solve this probability using the binomial distribution

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

The given values are: x = 1, n = 3, p = 0.466, $q = 1 - p$ = 0.534

$$P(X = 1) = \binom{3}{1}(0.466)^1(0.534)^2 = 0.399$$

You may also use the applet and should get the same result:

Binomial Distribution
$$X \sim Bin(n, p)$$

$n =$ `3` $p =$ `0.466`

$x =$ `1`   P(X = x) = ⌄ `0.39865`

$\mu = E(X) = 1.398 \qquad \sigma = SD(X) = 0.864 \qquad \sigma^2 = Var(X) = 0.747$

Help

6. What is the probability that the first two persons selected will be vaccine hesitant?

Your solution should be very similar to Question 4.

$$P(\text{only first two persons hesitant}) = P(\text{1st person hesitant}) \times P(\text{2nd person hesitant}) \times P(\text{3rd person not hesitant})$$
$$= 0.466 \times 0.466 \times 0.534$$
$$= 0.116$$

7. What is the probability that the exactly two persons selected will be vaccine hesitant?
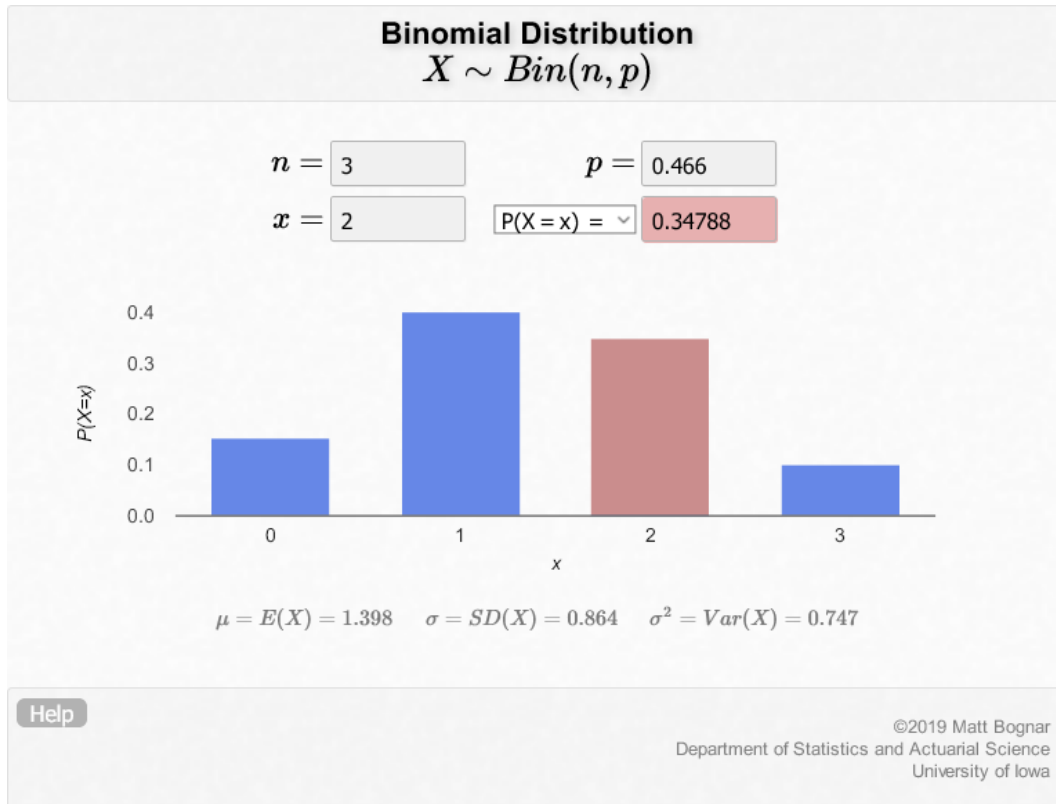
Your solution should be very similar to Question 5.

Using the addition rule:

$$P(\text{only two hesitant}) = P(\text{1st and 2nd hesitant}) + P(\text{1st and 3rd hesitant}) + P(\text{2nd and 3rd hesitant})$$
$$= 0.116 + 0.116 + 0.116$$
$$= 0.348$$

Using the binomial distribution formula, the given values are: x = 2, n = 3, p = 0.466, q = 0.534

$$P(X = 2) = \binom{3}{2}(0.466)^2(0.534)^1 = 0.348$$

You may also use the applet and should get the same result:

Binomial Distribution
$$X \sim Bin(n, p)$$

$n = $ 3 $p = $ 0.466

$x = $ 2 $P(X = x) = $ 0.34788

$\mu = E(X) = 1.398$ $\sigma = SD(X) = 0.864$ $\sigma^2 = Var(X) = 0.747$

Help

8. Using what you have calculated so far, and calculating some more, complete the following table:

Number of vaccine hesitant (x)	P(X = x) = Probability of x vaccine hesitant
0	
1	
2	
3	

Question 5 gave us $P(X = 1)$, while Question 7 gave us $P(X = 2)$. We can use either the formula or the applet to calculate $P(X = 0)$ and $P(X = 3)$. The applet is much easier to use, so let us plug in the values there.

Binomial Distribution
$$X \sim Bin(n, p)$$

$n =$ 3 $p =$ 0.466

$x =$ 3 P(X = x) = 0.10119
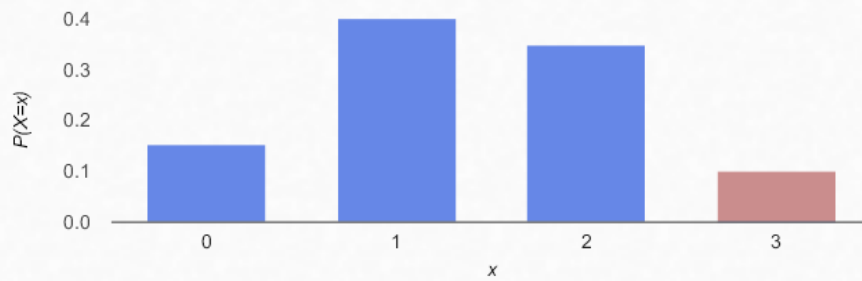
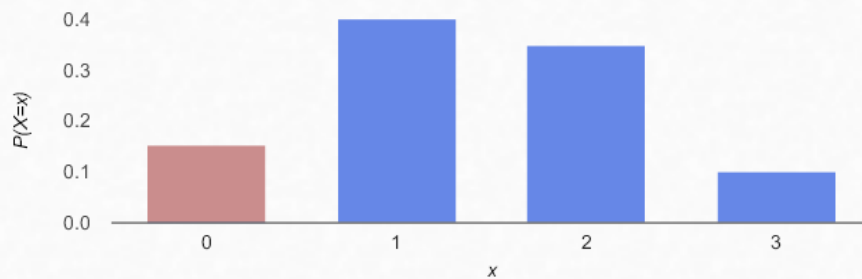$\mu = E(X) = 1.398$ $\sigma = SD(X) = 0.864$ $\sigma^2 = Var(X) = 0.747$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa

Binomial Distribution
$$X \sim Bin(n, p)$$

$n =$ 3 $p =$ 0.466

$x =$ 0 P(X = x) = 0.15227

$\mu = E(X) = 1.398$ $\sigma = SD(X) = 0.864$ $\sigma^2 = Var(X) = 0.747$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa
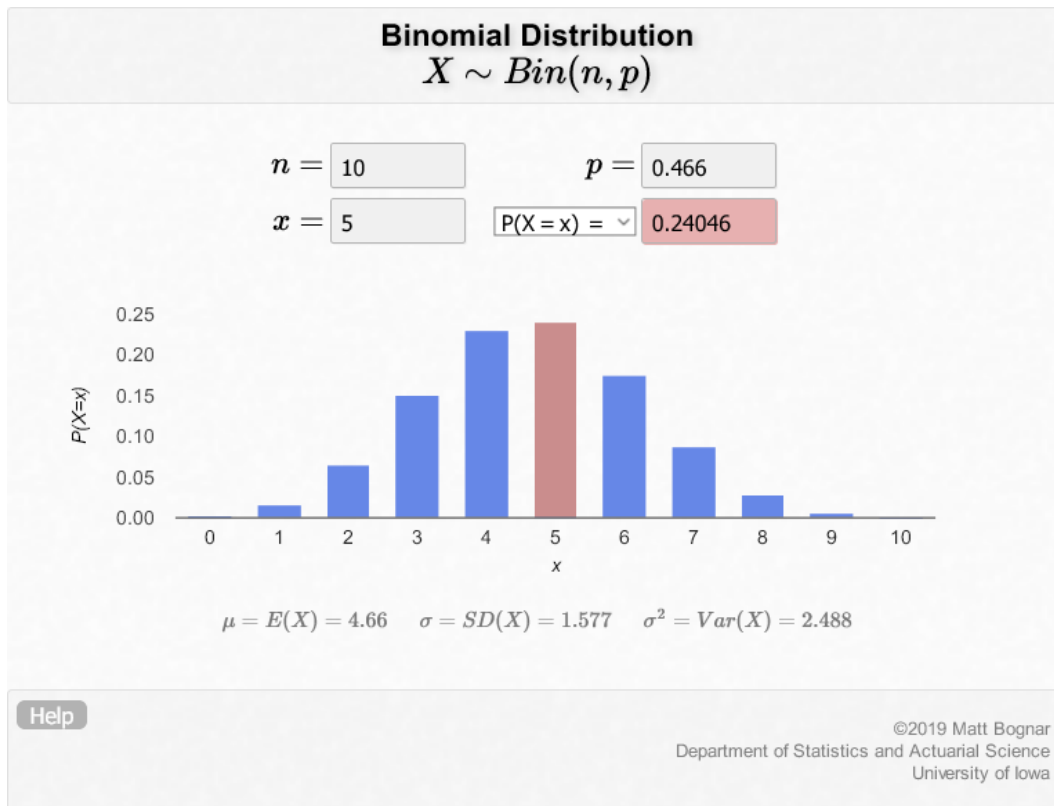
Therefore the values are:

Number of vaccine hesitant (x)	P(X = x) = Probability of x vaccine hesitant
0	0.153
1	0.348
2	0.399
3	0.101

Notice how these values correspond to the bars of each of the applet screenshots. Try hovering over the bars.

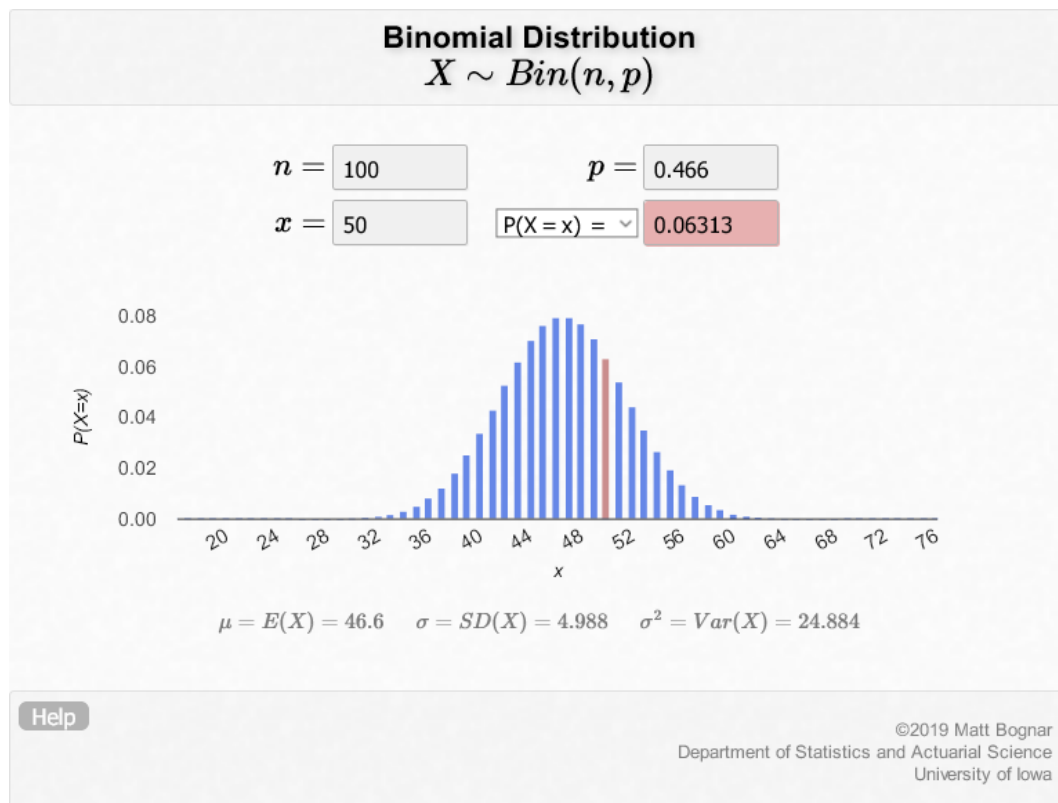**For questions 9 and 10, use the applet. Approximate to the nearest thousandths**

9. Now let's assume we are randomly selecting 10. What is the probability that 5 of them are going to be vaccine hesitant?

The probability is 0.240.

Binomial Distribution
$$X \sim Bin(n, p)$$

$n = 10$ $p = 0.466$

$x = 5$ P(X = x) = 0.24046

$\mu = E(X) = 4.66$ $\sigma = SD(X) = 1.577$ $\sigma^2 = Var(X) = 2.488$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa

10. Now let's assume we are randomly selecting 100. What is the probability that 50 of them are going to be vaccine hesitant? What do you notice about the binomial distribution as n becomes larger and larger?

The probability is 0.063. As n becomes larger, the binomial distribution resembles a normal distribution.

Binomial Distribution
$$X \sim Bin(n, p)$$

$n =$ `100` $p =$ `0.466`

$x =$ `50`   P(X = x) = `0.06313`

$\mu = E(X) = 46.6$ $\sigma = SD(X) = 4.988$ $\sigma^2 = Var(X) = 24.884$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa

PART B. Normal Distribution

In the same survey, individuals were asked about their perceptions of how other people will accept the COVID-19 vaccine, as a measurement of vaccine norms. They were asked, "Out of 100 people in your community, how many do you think would take a COVID-19 vaccine if it were made available?" We will call this variable **vaccine norms.** While this is technically a discrete variable, we will assume for the purposes of this class that this is a continuous variable.

In the latest survey wave, conducted March 14 to 28, 2021, the mean vaccine norm was 60.7 for the Philippines. We will assume that the population distribution is normal. We will also assume 60.7 is the true population distribution μ. We will also assume that the true population standard deviation σ is 12.5.

To answer the following questions, either calculate these by hand using a calculator or use the University of Iowa Normal Distribution Applet (https://homepage.divms.uiowa.edu/~mbognar/applets/normal.html). Approximate your answers to the nearest thousandth.
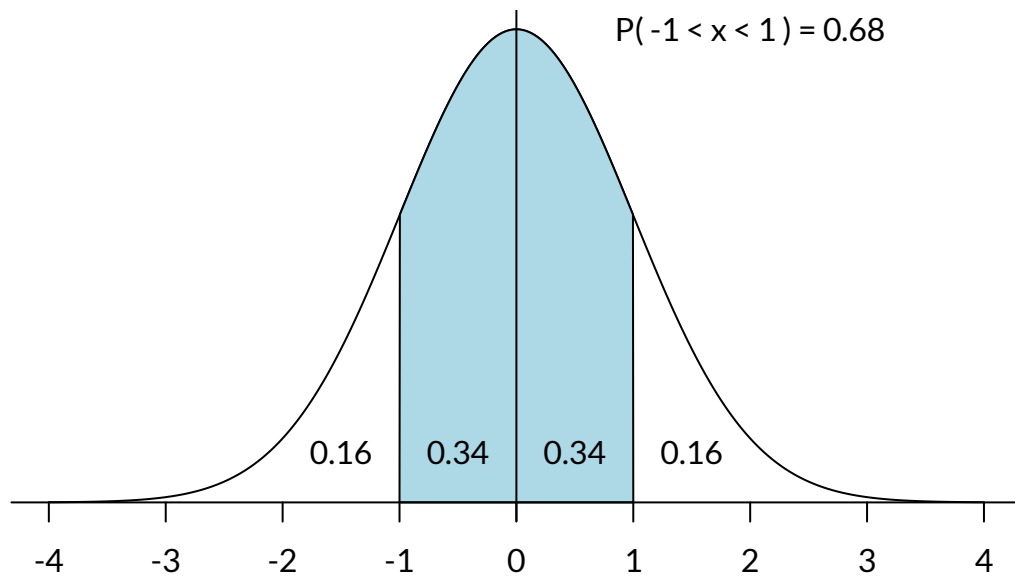
1. Approximately the middle 68% of Filipinos have vaccine norms of what range of values?

The question is asking for values approximately 1 standard deviation away from the mean or z = 1. Therefore, the range of values are from (60.7 - 12.5) to (60.7 + 12.5) = 48.2 to 73.2.

You may also use the applet to confirm these results.
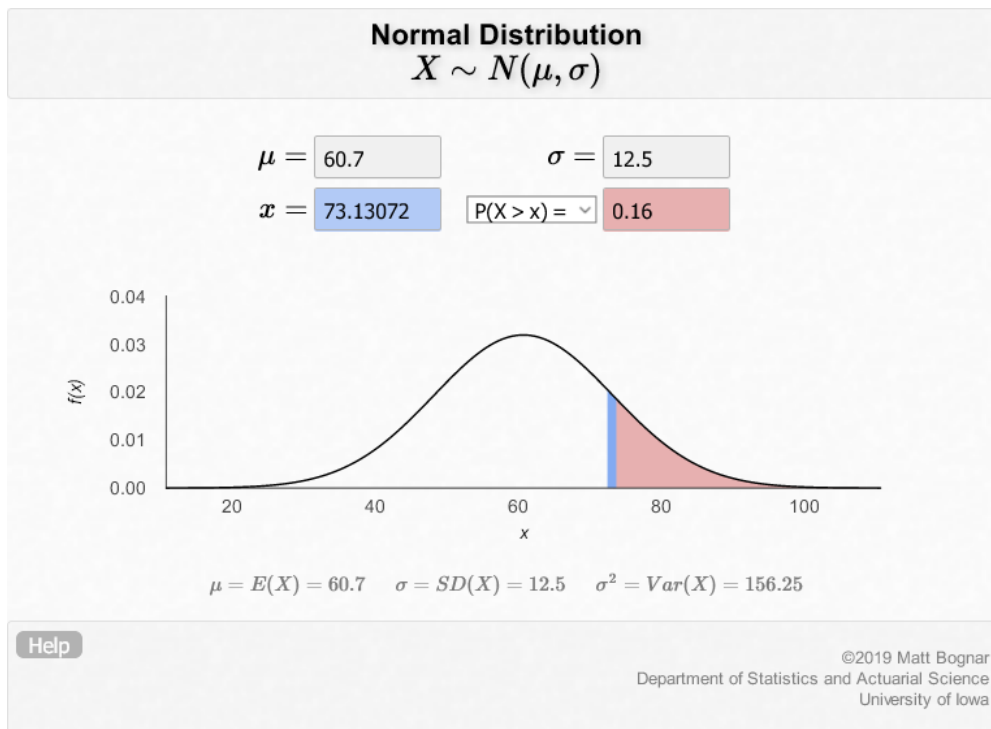
Based on the values given, μ = 60.7, σ = 12.5. Using the applet, we need to calculate the $P(X > x)$ values we need to get the corresponding x, which will define the upper and lower limits of the range of values. From the figure of the standard normal distribution below, This means that we need to get $P(X > x)$ = 0.16 and $P(X > x)$ = 0.16 + 0.34 + 0.34 = 0.84.

Standard Normal Distribution

P(-1 < x < 1) = 0.68

| 0.16 | 0.34 | 0.34 | 0.16 |

-4   -3   -2   -1   0   1   2   3   4

Number of standard deviations

Plugging in these values using the applet:

Normal Distribution
$$X \sim N(\mu, \sigma)$$

$\mu =$ 60.7 $\sigma =$ 12.5

$x =$ 73.13072   P(X > x) = 0.16

0.04
0.03
f(x) 0.02
0.01
0.00

20   40   60   80   100

x

$\mu = E(X) = 60.7$ $\sigma = SD(X) = 12.5$ $\sigma^2 = Var(X) = 156.25$

Help

Normal Distribution
$$X \sim N(\mu, \sigma)$$

$\mu =$ 60.7 $\sigma =$ 12.5

$x =$ 48.26928   P(X > x) = 0.84

$\mu = E(X) = 60.7$ $\sigma = SD(X) = 12.5$ $\sigma^2 = Var(X) = 156.25$

Help

Take a moment to absorb how the P(X > x) value works in the applet.

2. Approximately the middle 95% of Filipinos have vaccine norms of what range of values?

Based on the values given, μ = 60.7, σ = 12.5. The question is asking for values approximately 2 standard deviations away from the mean or *z* = 2. Therefore, the range of values are from (60.7 - 12.5 * 2) to (60.7 + 12.5 * 2) = 35.7 to 85.7.

You may also use the applet to confirm these results.

Based on the values given, μ = 60.7, σ = 12.5. Using the applet, we need to calculate the $P(X > x)$ values we need to get the corresponding x, which will define the upper and lower limits of the range of values. From the figure of the standard normal distribution below, This means that we need to get $P(X > x)$ = 0.025 and $P(X > x)$ = 0.025 + 0.475 + 0.475 = 0.975.
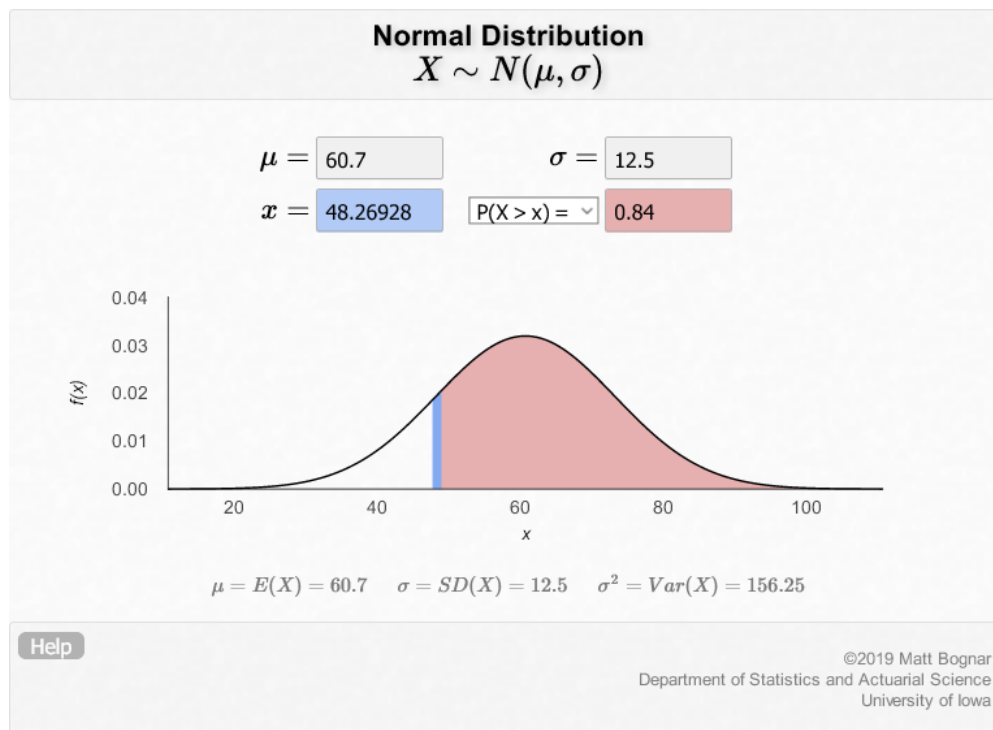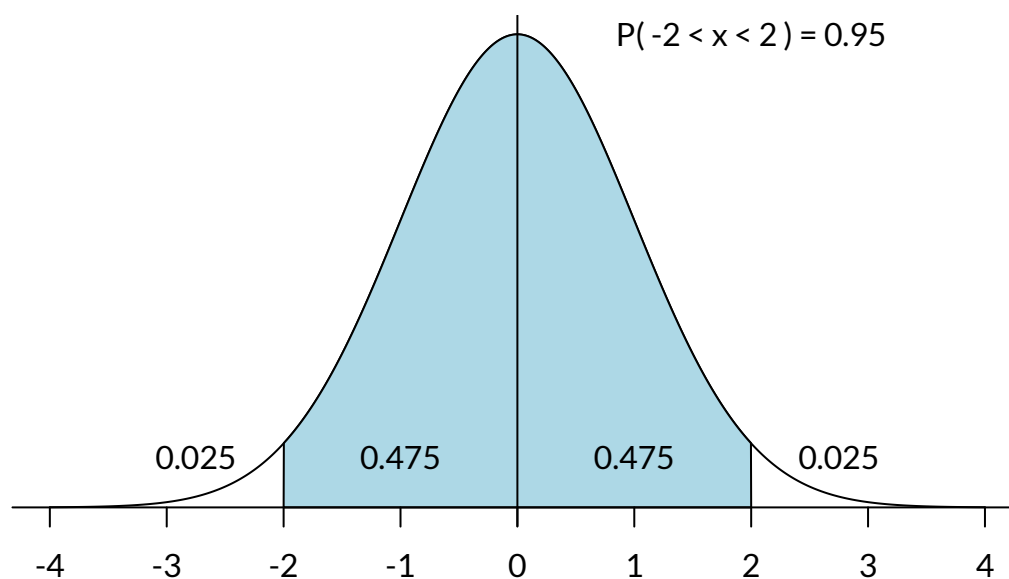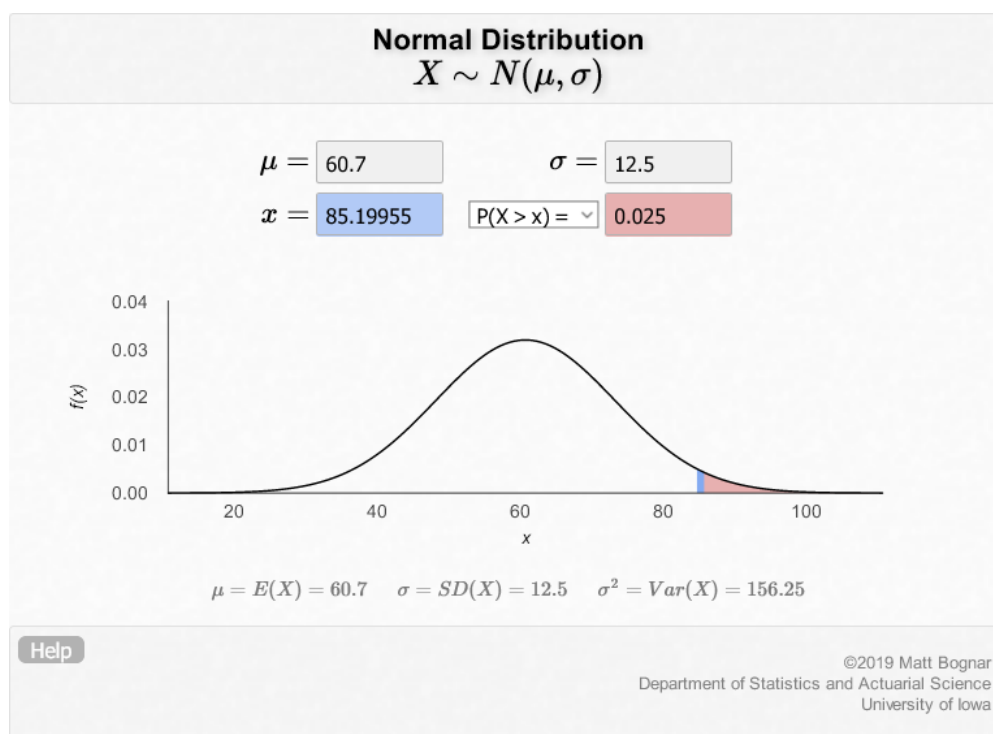
Standard Normal Distribution

P(-2 < x < 2) = 0.95

| 0.025 | 0.475 | 0.475 | 0.025 |

-4   -3   -2   -1   0   1   2   3   4

Number of standard deviations

Plugging in these values using the applet:

Normal Distribution
$$X \sim N(\mu, \sigma)$$

$\mu =$ 60.7 $\sigma =$ 12.5

$x =$ 85.19955   P(X > x) =   0.025

$\mu = E(X) = 60.7$ $\sigma = SD(X) = 12.5$ $\sigma^2 = Var(X) = 156.25$

Help

Normal Distribution
$X \sim N(\mu, \sigma)$

$\mu = $ 60.7 $\sigma = $ 12.5

$x = $ 36.20045   P(X > x) = 0.975

$\mu = E(X) = 60.7$ $\sigma = SD(X) = 12.5$ $\sigma^2 = Var(X) = 156.25$

Help
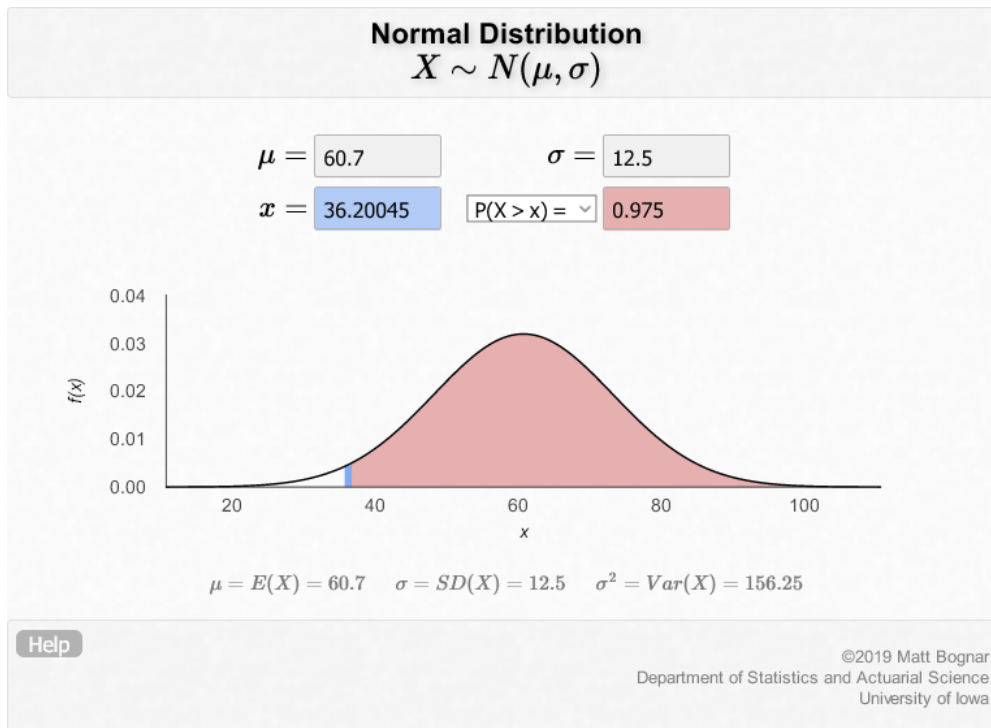
Note that the values in the applet and we calculated by hand are slightly different because we rounded off the *z* value to 2. To be more exact, the corresponding 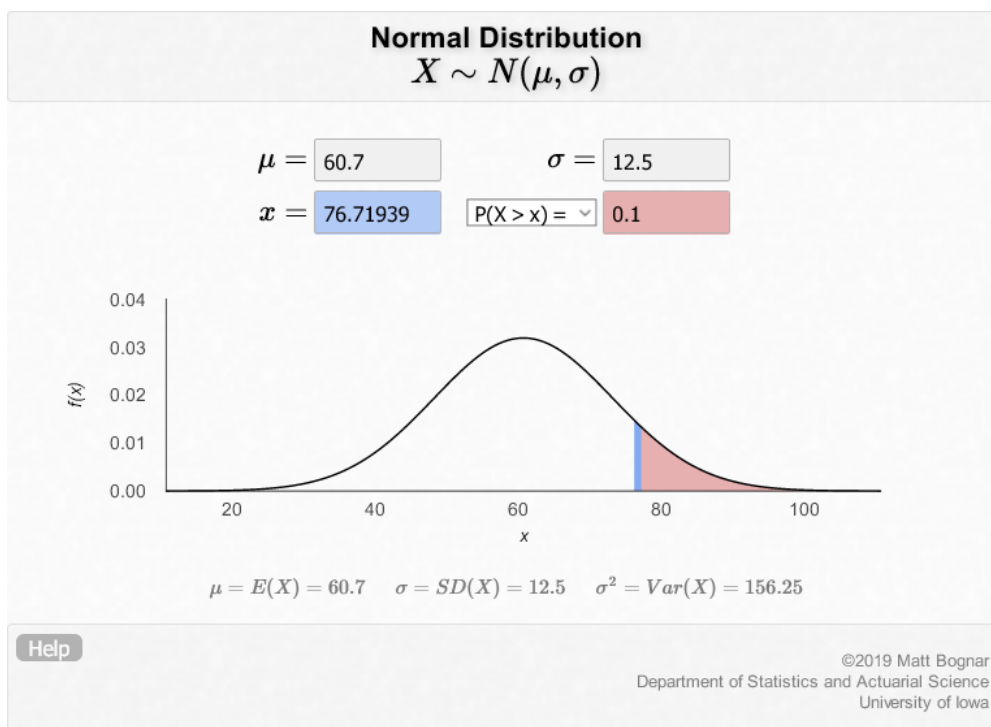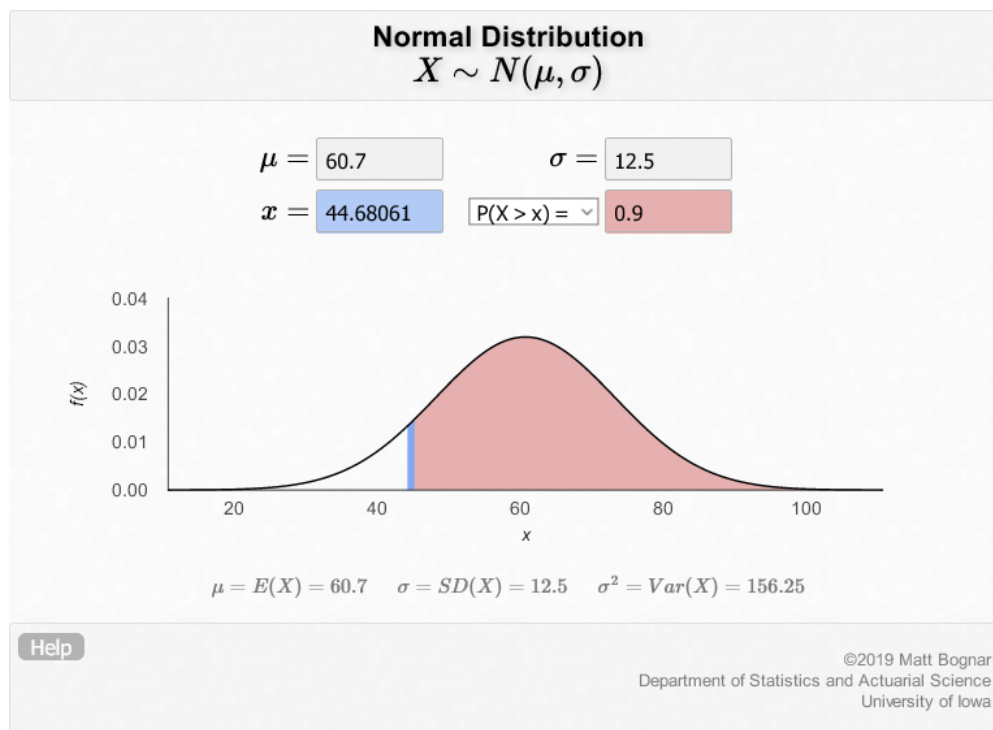*z* for 95% probability is 1.96. Therefore, the range of values are from (60.7 - 12.5 * 1.96) to (60.7 + 12.5 * 1.96) = 36.2 to 85.2, which is the same as the applet.

3. Approximately the middle 80% of Filipinos have vaccine norms of what range of values?

Using the applet, set $P(X > x)$ to 0.9 then 0.1. The range of values is 44.7 to 76.7

Normal Distribution
$X \sim N(\mu, \sigma)$

$\mu = $ 60.7 $\sigma = $ 12.5

$x = $ 76.71939   P(X > x) = 0.1

$\mu = E(X) = 60.7$ $\sigma = SD(X) = 12.5$ $\sigma^2 = Var(X) = 156.25$

Help

Normal Distribution
$$X \sim N(\mu, \sigma)$$

$\mu =$ 60.7 $\sigma =$ 12.5

$x =$ 44.68061   P(X > x) =   0.9

$\mu = E(X) = 60.7$ $\sigma = SD(X) = 12.5$ $\sigma^2 = Var(X) = 156.25$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa

4. How many percent of Filipinos have vaccine norms above 50?

Using the applet, set *x* = 50. The probability is 0.804

Normal Distribution
$$X \sim N(\mu, \sigma)$$

$\mu =$ 60.7 $\sigma =$ 12.5

$x =$ 50   P(X > x) =   0.804

$\mu = E(X) = 60.7$ $\sigma = SD(X) = 12.5$ $\sigma^2 = Var(X) = 156.25$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa
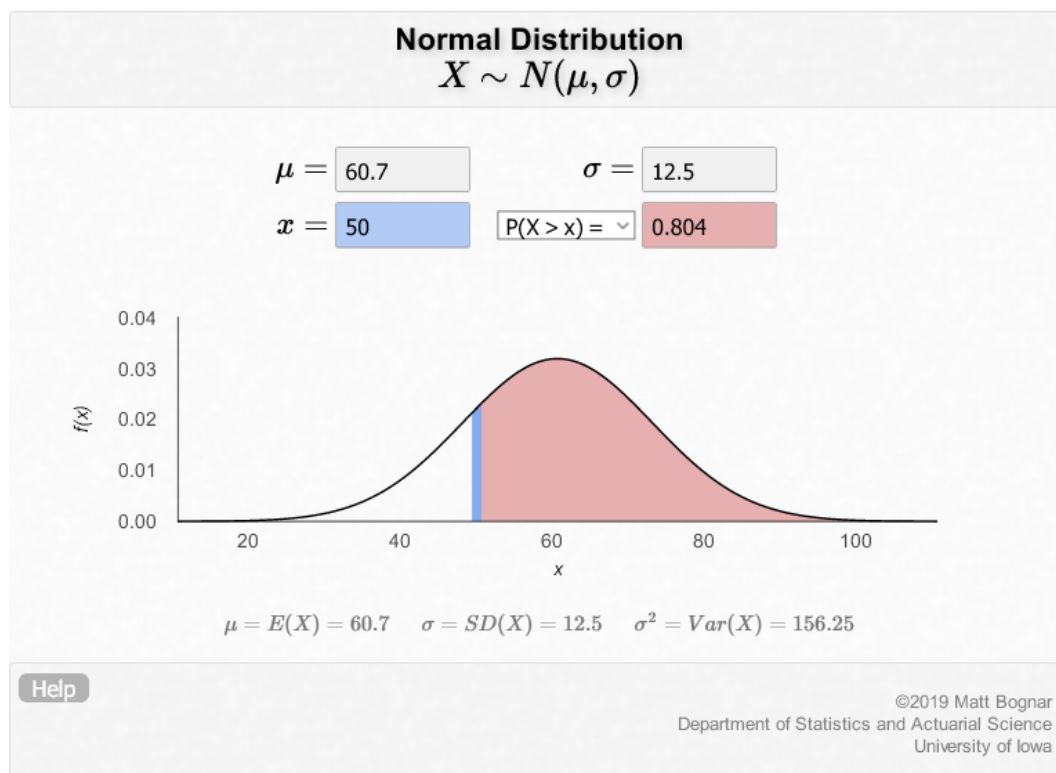
5. How many percent of Filipinos have vaccine norms between 35 and 55?

Change the setting of the applet from $P(X > x)$ to $P(X < x)$ since we need to subtract the values of two areas: $P(X < 55)$ and $P(X < 35)$. Set x = 55 then x = 35, then subtract the probabilities.

$$P(35 < X < 55) = P(X < 55) - P(X < 35) = 0.32419 - 0.01989 = 0.3043$$

Normal Distribution
$$X \sim N(\mu, \sigma)$$

$\mu =$ 60.7 $\sigma =$ 12.5

$x =$ 55   P(X < x) = ⌄ 0.32419

$\mu = E(X) = 60.7 \quad \sigma = SD(X) = 12.5 \quad \sigma^2 = Var(X) = 156.25$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa

Normal Distribution
$$X \sim N(\mu, \sigma)$$

$\mu =$ 60.7 $\sigma =$ 12.5

$x =$ 35   P(X < x) = ⌄ 0.01989

$\mu = E(X) = 60.7 \quad \sigma = SD(X) = 12.5 \quad \sigma^2 = Var(X) = 156.25$

Help

©2019 Matt Bognar
Department of Statistics and Actuarial Science
University of Iowa

END OF LAB