

Normal Distributions

Data seems to follow a predictable pattern where the majority of individuals are clumped in the center, and the greater the distance that a value is from the center, the fewer the individuals are that have that value. If someone was to draw a smooth curve around the picture that was just described you would end up with bell shaped curve.

To approximate the sampling distribution of \bar{x} as approximately normal see the guidelines on page 308 of Introductory Applied Statistics: A variable Approach.

A standard normal distribution is a normal distribution with a mean of 0 and a standard deviation of 1. A z-score tells us how many standard deviations away from the mean an observation is. We can use Z-scores to find certain probabilities.

Example Problem: The grades on a 215 statistics midterm at GVSU were normality distributed with $\mu = 85$ and a $\sigma = 2$.

a) If William scored a 74 on the midterm what was his z-score?

$$z = \frac{x - \mu}{\sigma} = \frac{74 - 85}{2} = -5.5$$

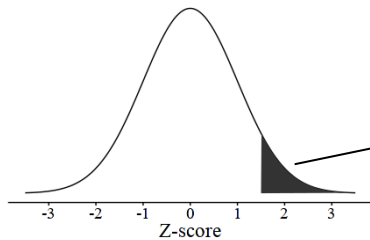
This tells us that William scored 5.5 standard deviations less than the mean.

b) What's the probability of scoring higher than 87?

Step 1: Find Z-score

$$z = \frac{x - \mu}{\sigma} = \frac{88 - 85}{2} = 1.5$$

Step 2: shade area under curve



The shaded area is the probability we are trying to find. The area under the curve is equal to 1.

Step 3: Use table to find probability or use calculator.

For this problem using the calculator:

$$\text{Normalcdf}(1.5, 1E99, 0, 1) = .1587$$

The lower bound is equal to 1.5 since our z-score is 1.5, and our upper bound is equal to infinity since we want to know the probability of scoring anything higher than an 87. Since we are using a z-score are mean is 0 and the stand deviation is 1.

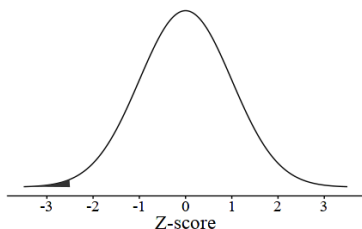
Calculator instructions for TI-84:
2nd VARS → Normalcdf(lower, upper, μ , σ)

c) What's the probability of scoring lower than an 80?

Step 1: Find Z-score

$$z = \frac{x - \mu}{\sigma} = \frac{80 - 85}{2} = -2.5$$

Step 2: shade area under curve



Step 3: Use table to find probability or use calculator.

Normalcdf(-1E99,-2.5,0,1) = .0067

Calculator instructions for TI-84:
2nd VARS → invnorm(probability,μ,σ)

d) If the probability of getting a score higher than X is .915, what is the value of X?

invnorm gives us the z-score based on the entered probability, mean (mu), and standard deviation (sigma).

Step 1: Find z-score

Invnorm(.915,0,1) = 1.3722

Step 2: Find Value of X using z-score formula

$$1.3722 = \frac{x-85}{2}$$

$$2.7444 = x - 85$$

$$x = \underline{87.744}$$

e) What is the probability of getting between a 75 and a 90 on the midterm?

Step 1: Find each values z-score

$$z = \frac{75-85}{2} = -5 \quad z = \frac{90-85}{2} = 2.5$$

Step 2: Find probability

Normalcdf(-5,2.5,0,1)=.9937

The lower bound is -5 because it's the z-score of the lowest value in the given range(70-90), 2.5 is the upper bound since it's the z-score of the highest value in the given range.

One Proportion Z-test

Example 2.19 on page 161 of Introductory Applied Statistics: A variable Approach. On November 4, 2008 American elected the first African-American in United States history to be President. An exit poll of 3,096 voters in 50 Michigan precincts was conducted that day by Edison Media Research and Mitofsky International. An exit poll is done by stopping voters as they leave the polling place and asking them a series of questions. Out of the 3096 voters polled, the number citing the economy as the most import issue was 2043. Did more than 50% of the voters polled think the economy was the most important issue?

Hypothesis: $H_0: p \leq 0.50$ $H_a: p > .50$

Conditions: $np_0 \geq 10$ $(3096)(0.5) = 1548$

$n(1 - p_0) \geq 10$ $(3096)(1-.05) = 1548$

Since the number of failures and the number of successes are each greater than 10 the conditions are met.

Test Statistic: $\hat{p} = \frac{2043}{3096}; p_0=0.50$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.660 - .50}{\sqrt{\frac{.50(1-.50)}{3096}}} = \frac{0.160}{0.008986} = 17.81$$

P-value: Can either use a table or a calculator. For a calculator use normalcdf(17.81,1E99,0,1)= 3.15×10^{-71}

Conclusion: Since our p-value is less than alpha=.05, we reject the null in favor of the alternative. There is sufficient evidence to conclude that the proportion of the voters polled think the economy was the most important issue is greater than 50%.