Problem Statement 1:

Blood glucose levels for obese patients have a mean of 100 with a standard deviation of 15. A researcher thinks that a diet high in raw cornstarch will have a positive effect on blood glucose levels. A sample of 36 patients who have tried the raw cornstarch diet have a mean glucose level of 108. Test the hypothesis that the raw cornstarch had an effect or not.

Solution:

Null Hypothesis: H0 = 100
Alternate Hypothesis: H1 ≠ 100

Standard Deviation : σ = 15

Average: μ = 100
Random sample size = n = 36
Xi = 108

Level of significance = 5% = 0.05= > If level of significance not given consider 5% significance
Two sided Test 0.05
Level of Confidence = 95%

The given sample size is greater than 30 and population is also known , hence consider
Z Statistic

Z = (x-μ)/ σ
Standard error SE(σ) = SD/√n
SE = 15 / √36 => 2.5
= (108-100)/2.5 => 3.2

Z Test = 3.2
P(Z<z) = 0.05

Check for 0.95 value in Z Table
Z(0.025) = 1.64
Calculated value 3.2 is in the Rejection region. Hence we reject accept Null Hypothesis.

Conclusion : we Reject the Null Hypothesis at 5% level of significance and it is concluded that there is raw cornstarch effect.

Problem Statement 2:

In one state, 52% of the voters are Republicans, and 48% are Democrats. In a second state, 47% of the voters are Republicans, and 53% are Democrats. Suppose a simple random sample of 100 voters are surveyed from each state.

What is the probability that the survey will show a greater percentage of Republican voters in the second state than in the first state?

Notes behind Solution:
Difference Between Proportions: Theory

Suppose we have two populations with proportions equal to P_1 and P_2. Suppose further that we take all possible samples of size n_1 and n_2. And finally, suppose that the following assumptions are valid.

- The size of each population is large relative to the sample drawn from the population. That is, N_1 is large relative to n_1, and N_2 is large relative to n_2. (In this context, populations are considered to be large if they are at least 10 times bigger than their sample.)
- The samples from each population are big enough to justify using a normal distribution to model differences between proportions. The sample sizes will be big enough when the following conditions are met: $n_1P_1 \geq 10$, $n_1(1 - P_1) \geq 10$, $n_2P_2 \geq 10$, and $n_2(1 - P_2) \geq 10$.
- The samples are independent; that is, observations in population 1 are not affected by observations in population 2, and vice versa.

Given these assumptions, we know the following.

- The set of differences between sample proportions will be normally distributed. We know this from the central limit theorem.
- The expected value of the difference between all possible sample proportions is equal to the difference between population proportions. Thus, $E(p_1 - p_2) = P_1 - P_2$.
- The standard deviation of the difference between sample proportions (σ_d) is approximately equal to:

$$\sigma_d = \text{sqrt}\{ [P_1(1 - P_1) / n_1] + [P_2(1 - P_2) / n_2] \}$$

It is straightforward to derive the last bullet point, based on material covered in previous lessons. The derivation starts with a recognition that the variance of the difference between independent random variables is equal to the sum of the individual variances. Thus,

$$\sigma^2_d = \sigma^2_{P1 - P2} = \sigma^2_1 + \sigma^2_2$$

If the populations N_1 and N_2 are both large relative to n_1 and n_2, respectively,

then $\sigma^2_1 = P_1(1 - P_1) / n_1$ And

$\sigma^2_2 = P_2(1 - P_2) / n_2$

Therefore,

$$\sigma^2_d = [P_1(1 - P_1) / n_1] + [P_2(1 - P_2) / n_2]$$

And

$\sigma_d = sqrt\{ [P_1(1 - P_1) / n_1] + [P_2(1 - P_2) / n_2] \}$

P_1 = the proportion of Republican voters in the first state
P_2 = the proportion of Republican voters in the second state
p_1 = the proportion of Republican voters in the sample from the first state
p_2 = the proportion of Republican voters in the sample from the second state
n1 = 100
n2 = 100

$n_1P_1 \geq 10$, $n_1(1 - P_1) \geq 10$, $n_2P_2 \geq 10$, and $n_2(1 - P_2) \geq 10$ ➔ Inorder to justify this :

- $n_1P_1 = 100 * 0.52 = 52$, $n_1(1 - P_1) = 100 * 0.48 = 48$, $n_2P_2 = 100 * 0.47 = 47$, and $n_2(1 - P_2) = 100 * 0.53 = 53$ are each greater than 10, the sample size is large enough

The expected value of the difference between all possible sample proportions is equal to the difference between population proportions. Thus, $E(p_1 - p_2) = P_1 - P_2$

- the mean of the difference in sample proportions: $E(p_1 - p_2) = P_1 - P_2 = 0.52 - 0.47 = 0.05$

Standard Deviation:

$$\sigma_d = sqrt\{ [P_1(1 - P_1) / n_1] + [P_2(1 - P_2) / n_2] \}$$

$\sigma_d = sqrt\{ [(0.52)(0.48) / 100] + [(0.47)(0.53) / 100] \}$

$\sigma_d = sqrt\ (0.002496 + 0.002491) = sqrt(0.004987) = 0.0706$

What is the probability that the survey will show a greater percentage of Republican voters in the second state than in the first state :

Inorder to calculate this use Z Test:

$$z_{p1 - p2} = (x - \mu_{p1 - p2}) / \sigma_d = = (0 - 0.05)/0.0706 = -0.7082$$

z-score being -0.7082
Probability = x < 0 : 0.2420

Conclusion :

Therefore, the probability that the survey will show a greater percentage of Republican voters in the second state than in the first state is 0.24.

Problem Statement 3:

You take the SAT and score 1100. The mean score for the SAT is 1026 and the standard deviation is 209. How well did you score on the test compared to the average test taker?

Solution:

$X = 1100$
$\mu = 1026$

Z Score = $(x - \mu)/\sigma$
Z score = $(1100 – 1026)/209$
= 0.354

This means that your score was 0.354 std devs above the mean.

How well did you score on the test compared to the average test taker

Z – value for 0.35 = 0.6368 = 63%