Z Tests and P-Values: Testing Hypotheses: σ is known and n > 30

Tests of the true value of an unknown population mean can be either one-tailed (left-tailed or right-tailed) or two-tailed.

1. two-tailed (non-directional)
   $H_0: \mu = \mu_0 \leftarrow$ a possibility you want to test (null hypothesis)
   $H_a: \mu \neq \mu_0 \leftarrow$ what the sample evidence suggests (alternative hypothesis)

<u>Reject H_0</u> if ($\bar{x}$ - μ_0) is a large positive number or a large negative number

2. left-tailed
   $H_0: \mu \geq \mu_0\}$ <u>Reject H_0</u> if ($\bar{x}$ - μ_0) is a large negative number.
   $H_a: \mu < \mu_0\}$

3. right-tailed
   $H_0: \mu \leq \mu_0\}$ <u>Reject H_0</u> if ($\bar{x}$ - μ_0) is a large positive number.
   $H_a: \mu > \mu_0\}$

Example: n = 50, σ = 4.6, $\bar{x}$ = 6.

1. $H_0: \mu = 5$ $H_a: \mu \neq 5$ $\alpha = .05$	2. $H_0: \mu \geq 6.5$ $H_a: \mu < 6.5$ $\alpha = .05$	3. $H_0: \mu \leq 4$ $H_a: \mu > 4$ $\alpha = .05$
$Z = \frac{6-5}{.65} = \frac{1}{.65} = 1.54$	$Z = \frac{6-6.5}{.65} = \frac{-.5}{.65} = -.77$	$Z = \frac{6-4}{.65} = \frac{2}{.65} = 3.08$

$\sigma / \sqrt{n} = 4.6/\sqrt{50}$

$\pm Z_{\alpha/2} = \pm 1.96$	$-Z_{\alpha} = -1.645$	$Z_{\alpha} = 1.645$
2-tailed rejection region at .05	**Both are 1-tailed rejection region at .05**	
Because Z = 1.54 is not <u>In the rejection area,</u> **<u>fail to reject H_0</u>**. The probability of being wrong in this conclusion is called p-value:	Because Z = -0.77 is <u>not in the rejection region,</u> **<u>fail to reject H_0</u>**.	Because Z = 3.08 <u>is in the rejection</u> region, **<u>reject H_0</u>**.
p = P(Z>1.54) + P(Z<-1.54) = .0618 x .0618	p = P (Z<-0.77) p = .2206 > .05	p = P (Z>3.08) p = .001 < .05

(**for 2-tailed test**) p = .1236 > .05

Remember 1: When looking-up the proportion in the tail in the Unit Normal Z table, the given p-value is for one-tailed tests. If you have a two-tailed test, as seen in example 1 on the previous page, multiply the given p-value by 2 to reflect the two-tailed nature of the test.

Remember 2: SPSS' p-values are presented as derived from two-tailed tests. If your alternative hypothesis being tested reflects a one-tailed test, you must divide the given SPSS p-value by 2 to reflect the one-tailed nature of your alternative hypothesis. From example 1 on the previous page, the p-value of .1236, reflecting a two-tailed test, would be readjusted by .1236 / 2 = one-tailed p-value of .0618.

Testing Hypotheses: σ is known and n > 30

I. One sample test for the population mean

1. H_0: $\mu = \mu_o$
2. H_a: $\mu \neq \mu_o$; $\mu < \mu_o$; $\mu > \mu_o$
3. Test statistic: Assume that H_0 is true and see if you have enough data/evidence to reject it.

3B. How far sample mean $\bar{x}$ is from μ

$$z = \frac{\bar{x} - \mu_o}{\sigma / \sqrt{n}}$$

4. P-value: Probability when H_0 is true of getting a test statistic as extreme as you did. Assume that H_0 is true until disproved.
5. Conclusion:
   A. Small p-value = reject H_0, there is enough evidence to say that H_a is true
      $\bar{x}$ is so far form μ_o that it is highly unlikely μ_o is true.
   B. Large p-value = fail to reject H_0. There is not sufficient evidence to say H_a is true.

1. Example:
The level of calcium in the blood of healthy, young adults varies with a mean of 9.5 mg per deciliter and a SD of 0.4. A clinic in rural Illinois measures the blood calcium level of 180 healthy pregnant women and finds $\bar{x}$ = 9.57mg. Is this an indication that the mean calcium level in this population differs from 9.5mg?

H_0: $\mu = 9.5$
H_a: $\mu \neq 9.5$

$$Z = \frac{\bar{x} - \mu_o}{\sigma / \sqrt{n}} \quad \frac{9.57 - 9.5}{.4 / \sqrt{180}} = \frac{.07}{.03} = 2.33$$

With alpha = .05 and because this is a two-tailed test (i.e., the = and the $\neq$), the critical region would consist of a Z score beyond ± 1.96 (note: this is found in the proportion of the tail, where .0250 is closest to Z = 1.96, so .0250 x 2 = .05).

Thus, with our Z value of 2.33 (look at table), the p-value is determined by .0099 (proportion of the tail section) x 2 (because of the two-tailed nature of this test) = .0198

There is only a 1.98% chance of getting a $\bar{x}$ of 9.57 or more extreme.

Conclusion:
At the .05 level, we would reject H_0 and say there is enough evidence to show the mean is different from 9.5. Thus, we have shown that the average level of calcium in the blood of pregnant Illinois women is different from 9.5 (or all other healthy young adults).

II. Example:

Mike gave the SAT math test to a simple random sample of 500 seniors from Illinois. These students had a mean score of 461 ($\bar{x}$). Is this good evidence that the mean for all Illinois seniors is > 450. $\sigma = 100$

H_0: $\mu \leq 450$

H_a: $\mu > 450$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{461 - 450}{100/\sqrt{500}} = \frac{11}{4.472} = 2.46$$

With alpha = .05 and because this is a one-tailed test (i.e., the ≤ and the >), the critical region would consist of a Z score beyond ± 1.645 (i.e., because a Z value of 1.64 and 1.65 are of equal distance in this case, we take the average).

Thus, with a Z value of 2.46, the p-value = .0069.

We reject H_0 at the .05 level and say that **there is a less than 1% chance of getting a $\bar{x}$ of 450 or more extreme.**

Conclusion:
A Z value of 2.46 indicates that our sample mean is in the critical region and this is a very unlikely outcome if H_0 is true and, thus, the decision to reject H_0. **The mean test score on the mathematics portion of the SAT for Illinois seniors is greater than 450.**