HYPOTHESIS TESTING: POWER OF THE TEST

The first 6 steps of the 9-step test of hypothesis are called **"the test"**. These steps are not dependent on the observed data values. When planning a research project, these steps are essential. However, one of the most important parts of your planning is determining how many observations you will need.

		H_o	
		true	false
What is the decision about H_o?	Reject H_o.	α	$1 - \beta$
	Accept H_o.	$1 - \alpha$	β

α = P(rejecting the null hypothesis given that it is true) = P(the observed value of the test statistic will fall in the rejection region when the null hypothesis is true).

β = P(accepting the null hypothesis given that it is false) = P(the observed value of the test statistic will not fall in the rejection region when the null hypothesis is false).

Power of the test = P(rejecting the null hypothesis given that it is not true) = P(the test statistic will fall in the rejection region when the null hypothesis is false). Power = $1 - \beta$.

Example

1. Given the following test, find β and the power of the test.

H_a: $\mu \neq 100$

H_o: $\mu = 100$

Assumptions: The random variable X is normally distributed with unknown mean μ and variance 64. A sample of size 16 is to be used.

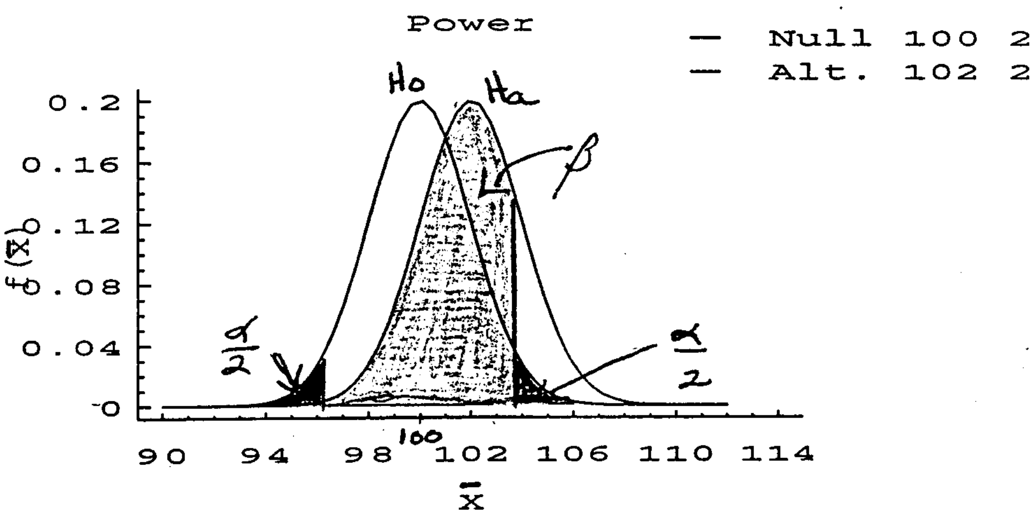
Test Statistic: $z = \dfrac{\overline{x} - \mu_o}{\dfrac{\sigma}{\sqrt{n}}}$

$\alpha = .05$ RR: $z < -1.96$ or $z > 1.96$

β = P(accepting the null hypothesis when it is false). To find this probability we need to express the critical values as values of the sample mean,

$\overline{x} < 100 - 1.96(8)/4 = 96.08$ or $\overline{x} > 100 + 1.96(8)/4 = 103.92$.

Thus, $\beta = P(\ 96.08 < \bar{x} < 103.92 \mid \mu_a \neq 100)$. Now we need the z-scores of 96.08 and 103.92 when μ_a is not 100. However, to do so we need to know which value of μ_a to use. The value of β and the power of the test will be functions of the values of μ_a, σ, n and α.



Let's compute β and the power of this test when $\mu_a = 102$. $\beta = P(\ 96.08 < \bar{x} < 103.92 \mid \mu_a = 102)$. Since $\sigma^2 = 64$ and n = 16, $\sigma_{\bar{x}} = 2$. Hence, $\beta = P(\ -2.96 < z < .96) = .8315 - .0015 = .8300$. Thus, power = $1 - \beta = 1 - .8300 = .1700$. Hence, if μ_a is actually 102 and we use a sample of size 16, there is only a 17% chance that we will reject the null hypothesis that $\mu = 100$.

The values of β and the power of this test for selected values of μ_a are given in the table. The plot of power versus μ_a is called the power curve for the test.

μ_a	β	Power
90	.0012	.9988
92	.0207	.9793
94	.1492	.8508
96	.4840	.5160
98	.8300	.1700
102	.8300	.1700
104	.4840	.5160
106	.1492	.8508
108	.0207	.9793
110	.0012	.9988

The formula for the power of a two-tailed test for the null hypothesis, $\mu = \mu_o$, is

$$Power = P\left(\bar{x} < -z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} + \mu_o \,\middle|\, \mu_a\right) + P\left(\bar{x} > z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} + \mu_o \,\middle|\, \mu_a\right) =$$

$$P\left(z < -z_{\frac{\alpha}{2}} - (\mu_a - \mu_o)\frac{\sqrt{n}}{\sigma}\right) + P\left(z > z_{\frac{\alpha}{2}} - (\mu_a - \mu_o)\frac{\sqrt{n}}{\sigma}\right)$$

This formula indicates that power is related to α, μ_a, σ and n.

What would happen if the test were a one-sided?

H_o: $\mu \leq 100$ versus H_a: $\mu > 100$

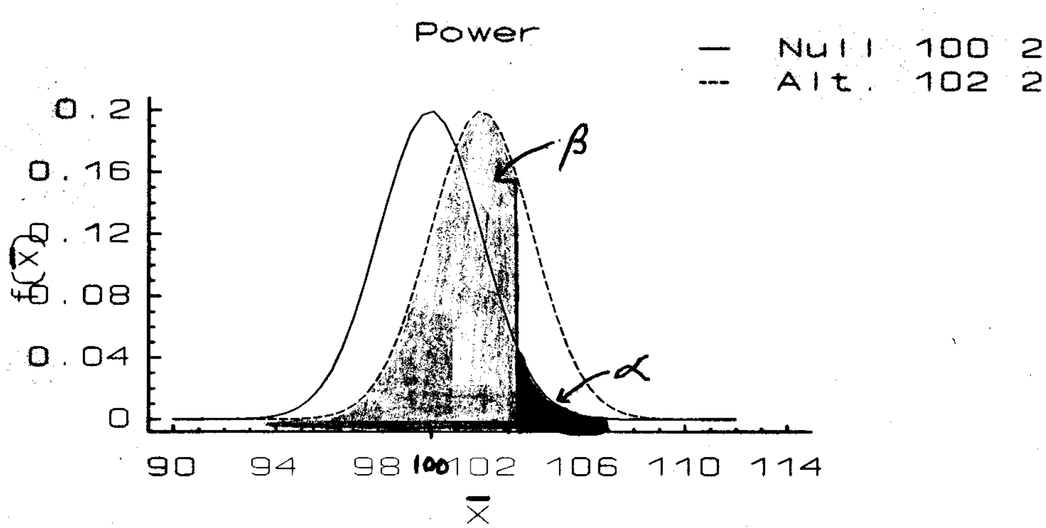
Test Statistic: $z = \dfrac{\bar{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$

$\alpha = .05$ RR: $z > 1.645$

To find β, we need to compute the probability of accepting the null hypothesis when it is not true. To find this probability we need to express the critical values as values of the sample mean,

$\bar{x} > 100 + 1.645(8)/4 = 103.29.$

Hence, $\beta = P(\bar{x} < 103.29 \mid \mu_a > 100)$. Once again we need the z-score of 103.29 when μ_a is greater than 100. To do this we need to know which μ_a to use. The values of β and power of the test will be a functions of the value of μ_a, α, σ and n.



Let's compute β and the power of this test when $\mu_a = 102$. $\beta = P(\bar{x} < 103.29 | \mu_a = 102)$. Since $\sigma^2 = 64$ and $n = 16$, $\sigma_{\bar{x}} = 2$. Hence, $\beta = P(z < .65) = .7422$. Thus, power $= 1 - \beta = 1 - .7422 = .2578$. Hence, if μ_a is actually 102 and we use a sample of size 16, there is only a 25.78% chance that we will reject the null hypothesis that $\mu = 100$.

The values of β and the power of this test for selected values of μ_a are given in the table.

μ_a	β	Power
102	.7422	.2578
104	.3594	.6406
106	.0869	.9131
108	.0091	.9909
110	.0004	.9996

There is a 25.78% chance of rejecting the null hypothesis that $\mu = 100$ when μ_a is actually 102 using a one-tailed test and a 17% chance when a two-tailed test is used. But when μ_a is actually 104 there is a 64.06% chance of rejecting the null hypothesis using a one-tailed test and a 51.60% chance using a two-tailed test.

When correctly, specified a one-tailed test is always more powerful than the corresponding two-tailed test.

The formula for the power of a one-tailed test for the null hypothesis, $\mu \leq \mu_o$, is

$$Power = P\left(\bar{x} > z_\alpha \frac{\sigma}{\sqrt{n}} + \mu_o \middle| \mu_a\right) = P\left(z > z_\alpha - (\mu_a - \mu_o)\frac{\sqrt{n}}{\sigma}\right).$$

This formula can be used to answer several different questions.

Examples: The test is

H_o: $\mu \leq 100$ versus H_a: $\mu > 100$

Assumptions: The population is normally distributed with $\sigma^2 = 64$ ($\sigma = 8$).

Test Statistic: $$z = \frac{\bar{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$$

$\alpha = .05$ RR: $z > 1.645$

1) Find the power of the test above when n = 20 and μ_a = 102.

Solution:

$$Power = P(\overline{x} > \frac{8(1.645)}{\sqrt{20}} + 100 | \mu_a = 102)$$

$$= P(z > 1.645 - \frac{(102-100)\sqrt{20}}{8}) = P(z > .53) = .2981$$

2) Find the sample size needed to have power = .83 when μ_a = 105 for the hypothesis test above.

Solution:

$$Power = .83 = P(\overline{x} > \frac{8(1.645)}{\sqrt{n}} + 100 | \mu_a = 105) =$$

$$P(z > 1.645 - \frac{(105-100)\sqrt{n}}{8}).$$

$$Let \;\; z_o = 1.645 - \frac{(105 - 100)\sqrt{n}}{8}, \;\; then \;\; P(z > z_o) = .83.$$

$$Therefore, \;\; z_o = -.95 = 1.645 - \frac{(105 - 100)\sqrt{n}}{8}.$$

$$Hence \;\; \sqrt{n} = 4.152 \; and \; n = 17.239104. \quad Use \;\; n = 18.$$

3) Find the value of μ_a when the power =.75 and n = 20 for the test above.

Solution:

$$Power = .75 = P(\overline{x} > \frac{8(1.645)}{\sqrt{20}} + 100 | \mu_a) = P(z > 1.645 - \frac{(\mu_a - 100)\sqrt{20}}{8})$$

$$Let \;\; z_o = 1.645 - \frac{(\mu_a - 100)\sqrt{20}}{8}, \;\; then \;\; P(z > z_o) = .75.$$

$$Therefore, \;\; z_o = -.67 = 1.645 - \frac{(\mu_a - 100)\sqrt{20}}{8}.$$

$$Hence \;\; \mu_a = \frac{(1.645 + .67)8}{\sqrt{20}} + 100 = 104.1411979.$$

Similar, computations are used to find the power for lower tailed tests.

If you examine each power formula, you find that each contain the expression $(\mu_a - \mu_o)\frac{\sqrt{n}}{\sigma}$. Let $\delta = (\mu_a - \mu_o)\frac{\sqrt{n}}{\sigma}$. We say that the effect size is **d** = $\frac{\mu_a - \mu_o}{\sigma}$. The general form of δ is given by $\delta = d[f(n)]$, where f(n) is a function of the sample size. For tests, such as matched pairs and two sample tests for the mean, see Howell's text for f(n). Howell's text also gives tables for finding the power of two-tailed tests based on α and δ.

EXERCISES

I. H_o: $\mu \leq 12$ versus H_a: $\mu > 12$

Assumptions: X is normally distributed with σ = 5.

Test Statistic: $z = \frac{\bar{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$

α = .01

1) For n = 14 and μ_a = 14, what is the power of this test? **DRAW A PICTURE.**

2) If the researcher wants the power of this test to be .85, how large should the sample size be? **DRAW A PICTURE.**

II. H_o: $\mu = 120$ versus H_a: $\mu \neq 120$

Assumptions: X is normally distributed with σ = 20.

Test Statistic: $z = \frac{\bar{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$

α = .05

1) For n = 25 and μ_a = 110, what is the power of this test? **Use both the formula and the table in the text.**

2) If the researcher wants the power of the test to be .90, how large should the sample size be when μ_a = 110? **Use both the formula and the table in the text.**

3) Redo parts 1 and 2 if α =;01.