STA258H5

Al Nosedal
and Alison Weir

Winter 2017

POWER AND SAMPLE SIZE

Power

The probability that a fixed level α significance test will reject H_0 when a particular alternative value of the parameter is true is called the **power** of the test against that alternative.

Sweetening colas: power

The cola maker of our example determines that a sweetness loss is too large to accept if the mean response for all tasters is $\mu = 1.1$. Will a 5% significance test of the hypotheses

$H_0 : \mu = 0$

$H_a : \mu > 0$

based on a sample of 10 tasters usually detect a change this great?

We want the power of the test against the alternative $\mu = 1.1$. This is the probability that the test rejects H_0 when $\mu = 1.1$ is true.

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}$.
We know that $\sigma = 1$, so the z test rejects H_0 at the $\alpha = 0.05$ level when

$$z = \frac{\bar{x} - 0}{1/\sqrt{10}} \geq 1.645$$

This is the same as

$$\bar{x} \geq 0 + 1.645 \frac{1}{\sqrt{10}}$$

or

$$\text{Reject } H_0 \text{ when } \bar{x} \geq 0.520$$

Step 2.

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 1.1$ is true.

To calculate this probability, standardize $\bar{x}$ using $\mu = 1.1$.

$$\begin{aligned} \text{power} &= P(\bar{x} \geq 0.520 \text{ when } \mu = 1.1) \\ &= P\left(\frac{\bar{x}-1.1}{1/\sqrt{10}} \geq \frac{0.520-1.1}{1/\sqrt{10}}\right) \\ &= P(Z \geq -1.83) = 1 - 0.0336 = 0.9664 \end{aligned}$$

Type I and Type II Errors

If we reject H_0 when in fact H_0 is true, this is a **Type I error.**
If we fail to reject H_0 when in fact H_a is true, this is a **Type II error.**
The **significance level** α of any fixed level test is the probability of a Type I error.
The **power** of a test against any alternative is 1 minus the probability of a Type II error for that alternative.

Example. Sweetening colas: error probabilities

When we select the significance level α of a test, we are setting the probability of a Type I error. Calculating the probability of a Type II error is just like calculating the power, except that we find the probability of the wrong decision (failing to reject H_0) rather than the probability of the right decision (rejecting).

$$H_0 : \mu = 0$$

$$H_a : \mu > 0$$

Example. Sweetening colas: error probabilities

The example shows that the z test rejects the null hypothesis at level $\alpha = 0.05$ when the mean sweetness loss assigned by 10 tasters satisfies $\bar{x} \geq 0.520$. The two error probabilities are

$$\begin{aligned} \text{P(Type I error)} &= \text{P(reject } H_0 \text{ when } \mu = 0) \\ &= P(\bar{x} \geq 0.520 \text{ when } \mu = 0) \\ &= P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \geq \frac{0.520-0}{1/\sqrt{10}}\right) \\ &= P(Z \geq 1.6443) = 0.05. \end{aligned}$$

Example. Sweetening colas: error probabilities

$$\begin{aligned}\text{P(Type II error)} &= \text{P(fail to reject } H_0 \text{ when } \mu = 1.1) \\ &= P(\bar{x} < 0.520 \text{ when } \mu = 1.1) \\ &= P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \frac{0.520-1.1}{1/\sqrt{10}}\right) \\ &= P(Z < -1.8341) = 0.0336.\end{aligned}$$

Example. Sweetening colas: error probabilities

$$
\begin{aligned}
\text{P(Type II error)} &= \text{P(fail to reject } H_0 \text{ when } \mu = 0.52) \\
&= P(\bar{x} < 0.520 \text{ when } \mu = 0.52) \\
&= P(\tfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \tfrac{0.520-0.52}{1/\sqrt{10}}) \\
&= P(Z < 0) = 0.5.
\end{aligned}
$$

R Code

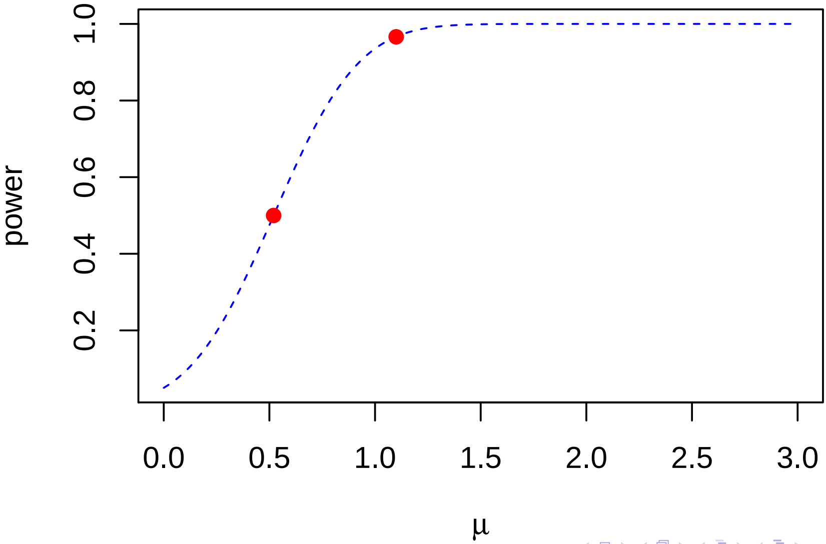
```
# Step 1. Computing power;

# vec.mu = vector with several different values of mu;
vec.mu = seq(0,3,by=0.01);
# n = sample size;
n = 10;
# sigma = population std. dev.;
sigma = 1;
# vec.z = vector of z-scores;
vec.z = sqrt(n)*(0.52-vec.mu)/sigma;
#beta = type II error;
beta = pnorm(vec.z);
power = 1 - beta;
```

R Code

```
# Step 2. Graphing power;

plot(vec.mu, power,type="l",lty=2,
col="blue", xlab=expression(mu));
points(0.52,0.5,pch=19,col="red");
points(1.1,1-0.0336,pch=19,col="red");
```


power
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
μ

Cola bottles: power

Bottles of a popular cola are supposed to contain 300 millimeters (ml) of cola. There is some variation from bottle to bottle because the filling machinery is not perfectly precise. The distribution of contents is Normal with standard deviation $\sigma = 3$ ml. Will inspecting 6 bottles discover underfilling? The hypotheses are

$$H_0 : \mu = 300$$

$$H_a : \mu < 300$$

Cola bottles: power

A 5% significance test rejects H_0 if $z_* \leq -1.645$, where the test statistic z_* is

$$z_* = \frac{\bar{x} - 300}{3/\sqrt{6}}$$

Power calculations help us see how large a shortfall in the bottle contents the test can be expected to detect. Find the power of this test against the alternative $\mu = 299$.

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}$.
We know that $\sigma = 1$, so the z test rejects H_0 at the $\alpha = 0.05$ level when

$$z = \frac{\bar{x} - 300}{3/\sqrt{6}} < -1.645$$

This is the same as

$$\bar{x} < 300 - 1.645\frac{3}{\sqrt{6}}$$

or

$$\text{Reject } H_0 \text{ when } \bar{x} < 297.985$$

Step 2.

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 299$ is true.

To calculate this probability, standardize $\bar{x}$ using $\mu = 299$.

$$\begin{aligned}\text{power} &= P(\bar{x} < 297.985 \text{ when } \mu = 299) \\ &= P\left(Z < \frac{297.985 - 299}{3/\sqrt{6}}\right) \\ &= P(Z < -0.83) = 0.2033\end{aligned}$$

Example

Suppose an experimenter wishes to test

$$H_0 : \mu = 100$$

$$H_a : \mu > 100$$

at the $\alpha = 0.05$ level of significance and wants $1 - \beta$ to equal 0.60 when $\mu = 103$. What is the smallest (i.e., cheapest) sample size that will achieve that objective? Assume that the variable being measured is Normally distributed with $\sigma = 14$.

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}_*$.
By definition,

$$\begin{aligned}\alpha &= \text{P(we reject } H_0 \text{ given } H_0 \text{ is true)}\\ &= P(\bar{X} > \bar{x}_* | \mu = 100)\\ &= P\left(\frac{\bar{X}-100}{14/\sqrt{n}} > \frac{\bar{x}_*-100}{14/\sqrt{n}}\right)\\ &= P\left(Z > \frac{\bar{x}_*-100}{14/\sqrt{n}}\right) = 0.05\end{aligned}$$

But $P(Z > 1.645) = 0.05$, so

$$\bar{x}_* = 100 + 1.645\frac{14}{\sqrt{n}}$$

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 103$ is true.
To calculate this probability, standardize $\bar{x}$ using $\mu = 103$.

$$\begin{aligned} \text{power } &= 1 - \beta = P(\bar{X} > \bar{x}_* \text{ when } \mu = 103) \\ &= P\left(\frac{\bar{X} - 103}{14/\sqrt{n}} > \frac{\bar{x}_* - 103}{14/\sqrt{n}}\right) = 0.60 \end{aligned}$$

From our Table, $P(Z > -0.25) = 0.5987 \approx 0.60$, so

$$\frac{\bar{x}_* - 103}{14/\sqrt{n}} = -0.25$$

which implies that $\bar{x}_* = 103 - 0.25\left(\frac{14}{\sqrt{n}}\right)$

Step 3. Solving for n

It follows from Steps 1 and 2 that

$$100 + 1.645\left(\frac{14}{\sqrt{n}}\right) = 103 - 0.25\left(\frac{14}{\sqrt{n}}\right)$$

(Solving for n)

$$n = \left[\frac{(1.645 + 0.25)(14)}{(103 - 100)}\right]^2 \approx 78.2045$$

Therefore, a minimum of 79 observations must be taken to guarantee that the hypothesis test will have the desired precision.

A vending machine advertises that it dispenses 225 ml cups of coffe ($\sigma = 7$ ml). You believe the mean volume of coffee per cup is something less than 225 ml. You plan to sample 40 cups of coffee from this machine to test your hypothesis.
a. If the true mean volume of coffee per cup is 223 ml, what is the power of your test at $\alpha = 0.05$? HW?
b. How many coffee cups should you sample if you want to raise the power in part (a) to 0.80?

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}_*$.

By definition,

$$\alpha = \text{P(we reject } H_0 \text{ given } H_0 \text{ is true)}$$
$$= P(\bar{X} < \bar{x}_* | \mu = 225)$$
$$= P\left(\frac{\bar{X}-225}{7/\sqrt{n}} < \frac{\bar{x}_*-225}{7/\sqrt{n}}\right)$$
$$= P\left(Z < \frac{\bar{x}_*-225}{7/\sqrt{n}}\right) = 0.05$$

But $P(Z < -1.645) = 0.05$, so

$$\bar{x}_* = 225 - 1.645\frac{7}{\sqrt{n}}$$

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 223$ is true.

To calculate this probability, standardize $\bar{x}$ using $\mu = 223$.

$$\text{power} = 1 - \beta = P(\bar{X} < \bar{x}_* \text{ when } \mu = 103)$$
$$= P\left(\frac{\bar{X} - 223}{7/\sqrt{n}} < \frac{\bar{x}_* - 223}{7/\sqrt{n}}\right) = 0.80$$

From our Table, $P(Z < 0.84) = 0.7995 \approx 0.80$, so

$$\frac{\bar{x}_* - 223}{7/\sqrt{n}} = 0.84$$

which implies that $\bar{x}_* = 223 + 0.84\left(\frac{7}{\sqrt{n}}\right)$

Solution b)

Step 3. Solving for n

It follows from Steps 1 and 2 that

$$225 - 1.645\left(\frac{7}{\sqrt{n}}\right) = 223 + 0.84\left(\frac{7}{\sqrt{n}}\right)$$

(Solving for n)

$$n = \left[\frac{(1.645 + 0.84)(7)}{(225 - 223)}\right]^2 \approx 75.6465$$

Therefore, a minimum of 76 observations must be taken to guarantee that the hypothesis test will have the desired precision.

Calculations of power (or of error probabilities) are useful for planning studies because we can make these calculations before we have any data. Once we actually have data, it is more common to report a P-value rather than a reject-or-not decision at a fixed significance level α. The P-value measures the strength of the evidence provided by the data against H_0. It leaves any action or decision based on that evidence up to each individual. Different people may require different strengths of evidence.