Creating decision rules and locating rejection regions

Example with a continuous random variable

Consider H_0: average weight of a male student in a certain college is 68kg. H_1: this average is not equal to 68. Formally:

Example with a continuous random variable

Consider H_0: average weight of a male student in a certain college is 68kg. H_1: this average is not equal to 68. Formally:

$$H_0 : \mu = 68$$

$$H_1 : \mu \neq 68$$

The alternative allows for the possibility of $\mu < 68$ or $\mu > 68$.

Evidence?

Example with a continuous random variable

Consider H_0: average weight of a male student in a certain college is 68kg. H_1: this average is not equal to 68. Formally:

$$H_0 : \mu = 68$$

$$H_1 : \mu \neq 68$$

The alternative allows for the possibility of $\mu < 68$ or $\mu > 68$.

Evidence?

A sample mean that falls close to 68 would be considered evidence in favor of H_0; but considerably less or more than 68 would be evidence against H_0.

Test statistic?

Example with a continuous random variable

Consider H_0: average weight of a male student in a certain college is 68kg. H_1: this average is not equal to 68. Formally:

$$H_0 : \mu = 68$$

$$H_1 : \mu \neq 68$$

The alternative allows for the possibility of $\mu < 68$ or $\mu > 68$.

Evidence?

A sample mean that falls close to 68 would be considered evidence in favor of H_0; but considerably less or more than 68 would be evidence against H_0.

Test statistic?

The sample mean.

What kind of data will lead you to reject H_0?

Here is a **proposal for a rejection region:** Reject H_0 if $\overline{X} < 67$ or $\overline{X} > 69$.

What kind of data will lead you to reject H_0?

Here is a **proposal for a rejection region:** Reject H_0 if $\overline{X} < 67$ or $\overline{X} > 69$.

Good? Bad? No idea?

- Arbitrary choice of the rejection region.
- Visualize the regions on a number line.

Example with a continuous random variable: solution

Assume that the standard deviation for the population of weights is $\sigma = 3.6$.

> *For large samples, we may substitute sample stdev (S) for σ, if not other estimate of σ is available.*

- Test statistic?

We will use $\overline{X}$, since this is a test about μ.

- Sampling distribution?

Sample size is $n = 36$. Central limit theorem $\implies$ distribution of $\overline{X}$ is approximately normal with $\sigma_{\overline{X}} = \frac{3.6}{6} = 0.6$.

- Decision rule / rejection region?

Reject $H_0 : \mu = 68$ if $\overline{X} < 67$ or $\overline{X} > 69$.

Plot: test in action

$$\textit{Reject } H_0 : \mu = 68 \textit{ if } \overline{X} < 67 \textit{ or } \overline{X} > 69.$$

Q: What is the probability of rejecting when H_0 is actually true?

$$P(\overline{X} < 67 \text{ or } \overline{X} > 69 \text{ when } H_0 \text{ is true}) = P(Z < a) + P(Z > b),$$

and we compute a, b as:

Plot: test in action

$$\textit{Reject } H_0 : \mu = 68 \textit{ if } \overline{X} < 67 \textit{ or } \overline{X} > 69.$$

Q: What is the probability of rejecting when H_0 is actually true?

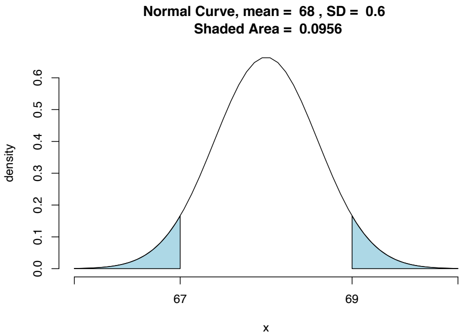
$$P(\overline{X} < 67 \text{ or } \overline{X} > 69 \text{ when } H_0 \text{ is true}) = P(Z < a) + P(Z > b),$$

and we compute a, b as: The **z-scores!**:

```
true.mu = 68
true.sigma = 3.6
n = 36
a= (67-68)/(true.sigma/sqrt(n))  # a -- the lower cut-off valu
b= (69-68)/(true.sigma/sqrt(n))  # b -- the upper cut-off valu
c(a,b)
```

```
[1] -1.666667  1.666667
```

```
pnormGC(bound=c(67,69), region="outside",
        mean=68, sd=true.sigma/sqrt(n),graph=TRUE)
```




```
[1] 0.0955807
```

The level of significance

Uh-oh... 9.5% **of all samples of size** 36 **would lead us to reject** $\mu = 68$ **kilograms when, in fact, it is true.**

Significance level

This error probability is called level of significance of the test, and denoted by α.

- Happy? ... seems too high of a chance of error.
- How to fix?
  - Increase the sample size (**try it yourself!**), or
  - Widen the fail-to-reject region.

The level of significance

Uh-oh... 9.5% **of all samples of size** 36 **would lead us to reject** $\mu = 68$ **kilograms when, in fact, it is true.**

Significance level

This error probability is called level of significance of the test, and denoted by α.

- Happy? ... seems too high of a chance of error.
- How to fix?
  - Increase the sample size (**try it yourself!**), or
  - Widen the fail-to-reject region.

But what about H_1?

Suppose H_1 is true and $\mu = 70$. What is $P(67 < \overline{X} < 69 \text{ when } \mu = 70)$?

Plot this!

Test significance levels. . . choices?

Philosophy: Preselection of significance level

Roots of pre-selection of α:

"The maximum risk of making a type I error should be controlled."

- Does not account for values of test statistics that are "close" to the critical region.
- Example: $H_0 : \mu = 10$ vs. $H_1 : \mu \neq 10$. Observed value $z = 1.87$.
  - with $\alpha = 0.05$, value not significant. (no reject)
  - but risk of error: $P = 2P(Z > 1.87 \text{ when } \mu = 10) = 2(0.0307) = 0.0614$.
  - 0.0614 is the probability of obtaining a value of z as large as or larger (in magnitude) than 1.87 when in fact $\mu = 10$.
  - $\implies$ Evidence against H_0 is not as strong as that which would result from rejection with $\alpha = 0.05$, but **it is important information to the user**.
  - Indeed, continued use of 'standard' $\alpha = 0.05$ or 0.01 only a result of what standards have been passed down through the generations.

Attained significance level

So how can we tell the user the important information about **strength of evidence**?

The *p*-value approach, adopted extensively by users of applied statistics, is designed to give the user an alternative (in terms of a probability) to a mere "reject" or "do not reject" conclusion.

- The P-value computation also gives the user important information when the z-value falls well into the ordinary critical region.
- For example, if $z = 2.73$, it is informative for the user to observe that $P = 2(0.0032) = 0.0064$, and thus the z-value is significant at a level considerably less than 0.05.
- It is important to know that under the condition of H_0, a value of $z = 2.73$ is an extremely rare event.
  - That is, a value at least that large in magnitude would only occur 64 times in $10,000$ experiments!

Graphical representation of *p*-value

Plot this!

Definition

A p-value is the lowest level (of significance) at which the observed value of the test statistic is significant.