M6S3 - Pvalue Interpretation

Professor Jarad Niemi

STAT 226 - Iowa State University

October 30, 2018

Outline

- Pvalues
    - Review of calculation procedure
    - Interpretation
- Hypothesis test
    - Decision making
    - Using p-values to make a decision
    - Errors
    -
- ASA Statement on P-values

P-values for $H_0 : \mu = m_0$

Definition

A p-value is the (frequency) probability of obtaining a test statistic as or more extreme than you observed if the null hypothesis (model) is true.

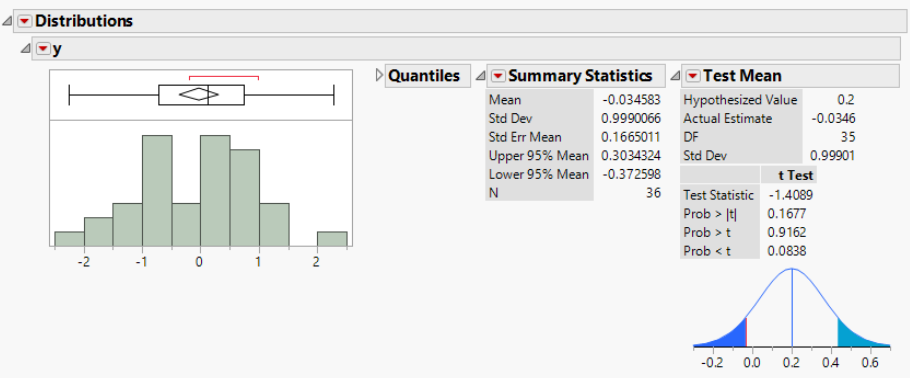
So for the null hypothesis $H_0 : \mu = m_0$, calculate

$$t = \frac{\overline{x} - m_0}{s/\sqrt{n}}$$

and find the appropriate probability:

- $H_a : \mu < m_0$ implies p-value $= \ \ P(T_{n-1} < t)$,
- $H_a : \mu > m_0$ implies p-value $= \ \ P(T_{n-1} > t)$, and
- $H_a : \mu \neq m_0$ implies p-value $= 2P(T_{n-1} > |t|)$,

JMP Example



Alternative Hypothesis	p-value
$H_a : \mu \neq m_0$	0.1677
$H_a : \mu > m_0$	0.9162
$H_a : \mu < m_0$	0.0838

Interpretation

Definition

A p-value is the probability of obtaining a test statistic as or more extreme than you observed if the null hypothesis is true.

If we assume $X_i \overset{iid}{\sim} N(m_0, \sigma^2)$ (because $H_0 : \mu = m_0$) and we have an observed test statistic t based on n observations, the p-value is

$$\begin{array}{ll} P(T_{n-1} < t) & \text{if } H_a : \mu < m_0 \\ P(T_{n-1} > t) & \text{if } H_a : \mu > m_0 \\ P(T_{n-1} > |t| \text{ or } T_{n-1} < -|t|) & \text{if } H_a : \mu \ne m_0 \end{array}$$

where

$$T_{n-1} = \frac{\overline{X} - m_0}{S/\sqrt{n}} \sim t_{n-1}$$

which is random because we are considering taking different random samples of size n.

Interpretation example

Assume $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ and $H_0 : \mu = 30$. From a random sample with 30 observations you find $\overline{x} = 35.2$ and $s = 11.57$ which results in

$$t = \frac{35.2 - 30}{11.57/\sqrt{30}} \approx 2.462.$$

You have the following probabilities

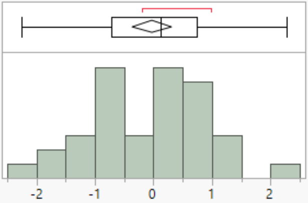
$$\begin{aligned}
P(T_{n-1} > 2.462) &= 0.01 \\
P(T_{n-1} < 2.462) &= 1 - P(T_{n-1} > 2.462) \\
&= 1 - 0.01 \\
&= 0.99 \\
P(T_{n-1} > 2.462 \text{ or } T_{n-1} < -2.462) &= P(T_{n-1} > 2.462) + P(T_{n-1} < -2.462) \\
&= 2P(T_{n-1} > 2.462) \\
&= 0.02
\end{aligned}$$

These probabilities correspond to the p-values for the alternative hypotheses $H_a : \mu > 30$, $H_a : \mu < 30$, and $H_a : \mu \neq 30$, respectively.

JMP Example

Distributions

y



Quantiles

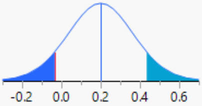
Summary Statistics

Mean	-0.034583
Std Dev	0.9990066
Std Err Mean	0.1665011
Upper 95% Mean	0.3034324
Lower 95% Mean	-0.372598
N	36

Test Mean

Hypothesized Value	0.2
Actual Estimate	-0.0346
DF	35
Std Dev	0.99901

	t Test
Test Statistic	-1.4089
Prob > |t|	0.1677
Prob > t	0.9162
Prob < t	0.0838

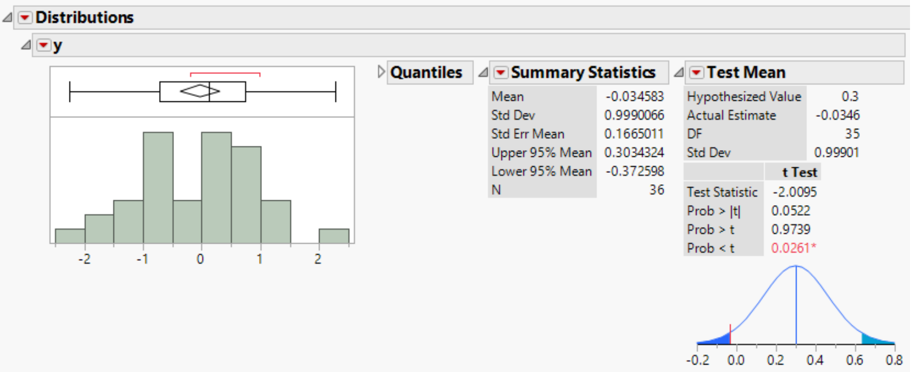


$$
\begin{aligned}
P(T_{35} > -1.4089) &= 0.9162 \\
P(T_{n-1} < -1.4089) &= 0.0838 \\
P(T_{n-1} > 1.4089 \text{ or } T_{n-1} < -1.4089) &= 0.1677
\end{aligned}
$$

Hypothesis test for a population mean μ

1. Specify the null and alternative hypothesis.
   - $H_0 : \mu = m_0$ is the default or current belief
   - $H_a : \mu > m_0$ or $\mu < m_0$ or $\mu \neq m_0$
2. Specify a significance level α.
3. Calculate the t-statistic.
4. Calculate the p-value.
5. Make a conclusion:
   - If p-value $< \alpha$, reject null hypothesis.
   - If p-value $\geq \alpha$, fail to reject null hypothesis.

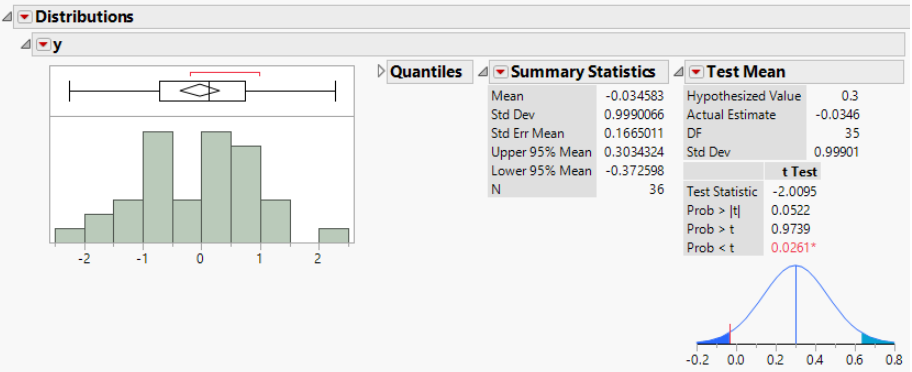
JMP Example



Conclusion at significance level $\alpha = 0.05$:

Alternative Hypothesis	p-value	Conclusion
$H_a : \mu \neq m_0$	0.0522	Fail to reject null hypothesis
$H_a : \mu > m_0$	0.9739	Fail to reject null hypothesis
$H_a : \mu < m_0$	0.0261	Reject null hypothesis

JMP Example



Conclusion at significance level $\alpha = 0.1$:

Alternative Hypothesis	p-value	Conclusion
$H_a : \mu \neq m_0$	0.0522	Reject null hypothesis
$H_a : \mu > m_0$	0.9739	Fail to reject null hypothesis
$H_a : \mu < m_0$	0.0261	Reject null hypothesis

Errors

When performing a hypothesis test, these are the possible situations:

	Truth	
Decision	H_0 true	H_0 not true
Reject H_0	Type I error	correct
Fail to reject H_0	correct	Type II error

Errors:

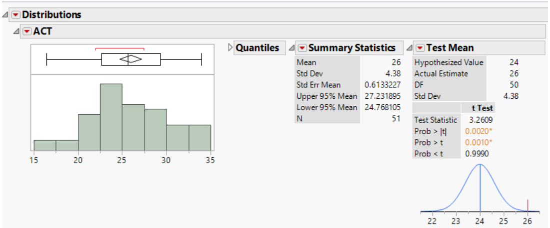
- A Type I error is rejecting the null hypothesis when it is true.
- A Type II error is failing to reject the null hypothesis when it is not true.

The significance level α is the Type I error.

ACT scores example

The mean composite score on the ACT among the students at Iowa State University is 24. We wish to know whether the average composite ACT score for business majors is different from the average for the University. Perform a hypothesis test at significance level $\alpha = 0.01$.

Let μ be the average mean composite score amount business majors at Iowa State University. Test $H_0 : \mu = 24$ versus $H_a : \mu \neq 24$.



Reject null hypothesis since $0.002 < 0.01$, i.e. p-value is less than significance level.

ASA Statement on p-values

In 2016, the American Statistical Association published Statement on p-values that states the following principles:

1. P-values can indicate how incompatible the data are with a specified statistical model.
2. P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
6. By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.

Interpretation (cont.)

The null hypothesis model is

$$X_i \stackrel{iid}{\sim} N(m_0, \sigma^2)$$

for some specificed value m_0.

P-values can indicate how incompatible the data are with [the null hypothesis] model.

The smaller the p-value the larger the incompatibility of the data with the null hypothesis model.

Thus, a small p-value indicates the null hypothesis model is likely not correct. But there are many assumptions in this model that may be wrong, e.g.

- independence,
- identically distributed,
- normality,
- mean is m_0, and
- constant variance.

The p-value doesn't tell us which one is wrong.