Lesson 033
Further Aspects of Hypothesis Testing

Friday, November 24

Practical Significance

- Hypothesis testing can tell us whether there is statistical evidence against a null hypothesis.
- It can not tell us whether or not we actually care.
- It is critical to differentiate **practical significance** from **statistical significance**.

Make sure to ask yourself:
"Even if this were true, would we care?"

Practical Significance

- The p-value is **not** a measure of the size of an effect.
  - With small sample sizes, even large effects may have large p-values.
  - With large sample sizes, even small effects may have small p-values.
- We cannot test for practical significance.

A hypothesis test considering the efficacy of a synthetic coating to the durability of a material is run, and the p-value is found to be very small ($p = 10^{-6}$). Which of the following is true?

The synthetic coating improves the performance substantially. 0%

The synthetic coating worsens the performance substantially. 0%

The synthetic coating improves the performance, but only mildly. 0%

The synthetic coating worsens the performance, but only mildly. 0%

None of the above are guaranteed to be true. 0%

Suppose a hypothesis test comparing the mean failure time of a particular component compared to a standard. Differences over 1000 hours have a large effect. The observed difference was $\hat{\Delta} = 5000$ with $p = 0.33$ for $H_0 : \Delta = 0$.

These results suggest a large, statistically significant effect.

0%

These results suggest a small, statistically significant effect.

0%

These results are statistically significant but not practically significant.

0%

There is likely no meaningful difference in mean failure time.

0%

Which of the following would likely suggest the most meaningful effects, if observed from a hypothesis test?

Effect size of 5 with a p-value of 0.01.

0%

Effect size of 10 with a p-value of 0.15.

0%

Effect size of 10 with a p-value of 0.01.

0%

Effect size of 0.01 with a p-value of 0.005.

0%

Confidence Intervals and Hypothesis Tests

- Every confidence interval can be used to generate hypothesis tests based on an α level of significance.
- If $H_0 : \theta = \theta_0$ then if the $(1 - \alpha)100\,\%$ CI contains the value θ_0, we fail to reject H_0 at an α level.
- Why?

A 95% confidence interval for μ is formed as $[-2, 3]$. Suppose that we wish to test $H_0 : \mu = -1$. Which of the following conclusions hold?

We reject H_0 at a 5% level of significance.

0%

We fail to reject H_0 at a 5% level of significance.

0%

We reject H_0 at a 1% level of significance.

0%

We fail to reject H_0 at a 1% level of significance.

0%

More than one of the above.

0%

A 95% confidence interval for μ is formed as $[-2, 3]$. Suppose that we wish to test $H_0 : \mu = 4$. Which of the following conclusions hold?

We reject H_0 at a 5% level of significance.

0%

We fail to reject H_0 at a 5% level of significance.

0%

We reject H_0 at a 10% level of significance.

0%

We fail to reject H_0 at a 10% level of significance.

0%

More than one of the above.

0%

Suppose that the hypothesis test of $H_0 : \mu = 0$ has a compute p-value of 0.03. Which of the following confidence intervals are plausible?

A 95% confidence interval of $[-1, 1]$.

0%

A 90% confidence interval of $[-1, 1]$.

0%

A 99% confidence interval of $[-1, 1]$.

0%

More than one of the above.

0%

Critiques of Hypothesis Testing

- Some statisticians deeply dislike hypothesis testing as a framework.
  - P-values are prioritized above all else.
  - The null hypothesis is often inadequate.
  - The framework itself is easily abused.
  - Even when it works, it can lead to concerns.

Suppose that the p-value for a (symmetric) test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ is 0.04. What would the p-value be for the same procedure of $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$.

0.02

0%

0.04

0%

0.96

0%

0.98

0%

We need more information.

0%

Suppose that the p-value for a (symmetric) test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ is 0.04. What would the p-value be for the same procedure of $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$. Suppose $\widehat{\mu} > \mu_0$.

0.02

0%

0.04

0%

0.96

0%

0.98

0%

We need more information.

0%

Suppose that the p-value for a (symmetric) test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ is 0.04. What would the p-value be for the same procedure of $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$. Suppose $\widehat{\mu} < \mu_0$.

0.02	0%
0.04	0%
0.96	0%
0.98	0%
We need more information.	0%

Suppose that $H_0 : \mu \geq \mu_0$ is tested against $H_1 : \mu < \mu_0$, where $\hat{\mu} < \mu_0$. The p-value is computed as 0.03. What would be the conclusion of testing $H_0 : \mu = \mu_0$ at $\alpha = 0.05$?

We fail to reject H_0.

0%

We reject H_0.

0%

We cannot test this from the given information.

0%