STATISTICS FOR DATA SCIENCE

HYPOTHESIS and INFERENCE

Dr. Deepa Nair

Department of Science and Humanities

STATISTICS FOR DATA SCIENCE

UNIT-4 HYPOTHESIS and INFERENCE
Session-4
Drawing Conclusions from the Results of Hypothesis Tests

Dr. Deepa Nair
Department of Science and Humanities

Drawing Conclusions from the Results of Hypothesis Tests

- The only two conclusions that can be reached in a hypothesis test are that

- H_0 is false or that H_0 is plausible.

- One can never conclude that H_0 is true.

Source: Text Book Chapter6

Drawing Conclusions from the Results of Hypothesis Tests

- **How do we know when to reject H_0?**
- **The smaller the P-value, the less plausible H_0 becomes.**
- **A common rule of thumb is to draw the line at 5%. According to this rule of thumb, if $P \leq 0.05$, H_0 is rejected; otherwise H_0 is not rejected.**

Source: Text Book Chapter6

Drawing Conclusions from the Results of Hypothesis Tests

- The smaller the P-value, the more certain we can be that H_0 is false.

- The larger the P-value, the more plausible H_0 becomes, but we can never be certain that H_0 is true.

Source: Text Book Chapter6

Drawing Conclusions from the Results of Hypothesis Tests

- There is no sharp dividing line between conclusive evidence against H_0

- So while this rule of thumb is convenient, it has no real scientific justification.

Source: Text Book Chapter6

- A rule of thumb suggests to reject H_0 whenever $P \leq 0.05$.

Statistical Significance:

- **Whenever the *P*-value is less than a particular threshold, the result is said to be "statistically significant" at that level.**

- **So, for example, if $P \le 0.05$, the result is statistically significant at the 5% level; if $P \le 0.01$, the result is statistically significant at the 1% level, and so on.**

Source: Text Book Chapter6

Statistical Significance:

- If a result is statistically significant at the $100\alpha\%$ level, we can also say that the null hypothesis is "rejected at level $100\alpha\%$."

Source: Text Book Chapter6

Drawing Conclusions from the Results of Hypothesis Tests

- **The null hypothesis is rejected at the $100\alpha\%$ level.**

- **When reporting the result of a hypothesis test, report the P −value, rather than just comparing it to 5% or 1%.**

Source: Text Book Chapter6

Drawing Conclusions from the Results of Hypothesis Tests

- **Let α be any value between 0 and 1. Then, if $P \leq \alpha$,**
- **The result of the test is said to be statistically significant at the $100\alpha\%$ level.**

Source: Text Book Chapter6

Drawing Conclusions from the Results of Hypothesis Tests

PES UNIVERSITY ONLINE

The P-value Is Not the Probability That H_0 Is True

- Since the $P-value$ is a probability, and since small $P-values$ indicate that H_0 is unlikely to be true, it is tempting to think that the $P-value$ represents the probability that H_0 is true. This is emphatically not the case. The concept of probability discussed here is useful only when applied to outcomes that can turn out in different ways when experiments are repeated. It makes sense to define the $P-value$ as the probability of observing an extreme value of a statistic such as X, since the value of X could come out differently if the experiment were repeated. The null hypothesis, on the other hand, either is true or is not true. The truth or falsehood of H_0 cannot be changed by repeating the experiment. It is therefore not correct to discuss the "probability" that H_0 is true.

Source: Text Book Chapter6

Drawing Conclusions from the Results of Hypothesis Tests

Example:

- **A hypothesis test is performed of the null hypothesis $H_0: \mu = 0$. The $P-$value turns out to be 0.03.**
- **Is the result statistically significant at the 10% level? The 5% level? The 1% level?**
- **Is the null hypothesis rejected at the 10% level? The 5% level? The1% level?**

Source: Text Book Chapter6

Solution:

- **The result is statistically significant at any level greater than or equal to 3%.**

- **Thus it is statistically significant at the 10% and 5% levels, but not at the 1% level.**

- **Similarly, we can reject the null hypothesis at any level greater than or equal to 3%**

- **So H_0 is rejected at the 10% and 5% levels, but not at the 1% level.**

Example:

- The length of life X of certain computers is approximately normally distributed with mean 800 hours and standard deviation 40 hours.

- If a random sample of 30 computers has an average life of 788 hours, test the null hypothesis that $\mu = 800\ hours$ against the alternate hypothesis that $\mu \neq 800\ hours$. at

- $\mathbf{0.5\%}\ \ b)\ \mathbf{1\%}\ \ c)\mathbf{4\%}\ \ d)\mathbf{5\%}\ \ e)\ \mathbf{10\%}\ \ f)\mathbf{15\%}$ level of significance.

Source: Higher Engineering Mathematics by B.V Ramana

Solution:

$$H_0:\ \mu = 800\ hours \quad ,H_1:\ \mu \neq 800\ hours$$

$$\bar{X} = 788, n = 30, \mu = 800, \sigma = 40$$

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{788 - 800}{40/\sqrt{30}} = -1.643$$

Solution:

Case 1:

$$\alpha = 0.5\% = 0.005$$

P- Value $0.101 > 0.005$

So we need to reject the null hypothesis at 0.5% level.

Solution:

Case 2:

$$\boldsymbol{\alpha = 1\% = 0.01}$$

P- Value $\mathbf{0.101 > 0.01}$

So we need to reject the null hypothesis at 1% level.

Solution:

Case 3:

$\alpha = 4\% = 0.04$

P- Value $0.101 > 0.04$

So we need to reject the null hypothesis at 4% level.

Solution:

Case 4:

$$\boldsymbol{\alpha = 5\% = 0.05}$$

P- Value $0.101 > 0.05$

So we need to reject the null hypothesis at 5% level.

Solution:

Case 5:

$\alpha = 10\% = 0.10$

P- Value $0.101 > 0.10$

So we need to reject the null hypothesis at 10% level.

Solution:

Case 6:

$\alpha = 15\% = 0.15$

P- Value $0.101 < 0.15$

So we accept the null hypothesis at 15% level.

Example:

- Mice with an average life span of 32months will live up to 40 months when fed by a certain nutrious food.
- If 64 mice fed on this diet have an average life span of 38 months and standard deviation of 5.8 months.
- Is there any reason to believe that the average life span is less than 40 months.

Solution:

Let us take 0.01 as the significance level.

$$H_0:\ \mu \geq 40\ months, \qquad H_1:\ \mu < 40\ months$$

$$\bar{X} = 38, \qquad n = 64, \qquad \sigma \to s = 5.8$$

$$z = \frac{38 - 40}{5.8/\sqrt{64}} = -2.76$$

P- Value is 0.0029 $0.0029 < 0.01$

We need to reject H_0

We will conclude that there is a reason to believe that the average life span of mice with nutrition food is less than 40 months

Dr. Deepa Nair
Department of Science and Humanities

deepanair@pes.edu