Confidence Intervals & Hypothesis Testing

Module 4d Conducting a Hypothesis Test

Learning Objectives

- How to conduct two-tailed and one-tailed hypothesis tests
- Understand why different calculations are required to conduct a hypothesis test about p than μ
- Using ISLE to conduct a hypothesis test

The Concept of α in Hypothesis Testing

- The Null Hypothesis, H_0 , is the favored hypothesis and the test is conducted under the assumption that it is true.
- Only reject H_0 in favor of the Alternative Hypothesis, H_1, if the sample strongly favors it.
- α called the significance level is the probability standard for making that determination.
- The most commonly used standard is $\alpha = 0.05$
- By this standard we only reject H_0 if there is only 5% chance or less that the sample mean could have been drawn if H_0 is true.
- The 5% standard was suggested by Sir Ronald Fisher in his 1935 book

Two-Sided Hypothesis Tests

We want to test H_1: $\mu \neq \mu_0$ against H_0: $\mu = \mu_0$.

<u>Solution</u>: reject the null hypothesis if $\bar{x}$ is sufficiently far from μ_0.

Assume $\mu = \mu_0$. According to the CLT, $\bar{x}$ is normally distributed with mean μ_0 and standard deviation $\sigma/\sqrt{N} \approx s/\sqrt{N}$. Thus the t-score of $\bar{x}$ is t = $(\bar{x} - \mu_0)$ / $s/\sqrt{N}$

Reject H_0 if t< $-t_{\alpha/2}^{df}$ or z > $t_{\alpha/2}^{df}$

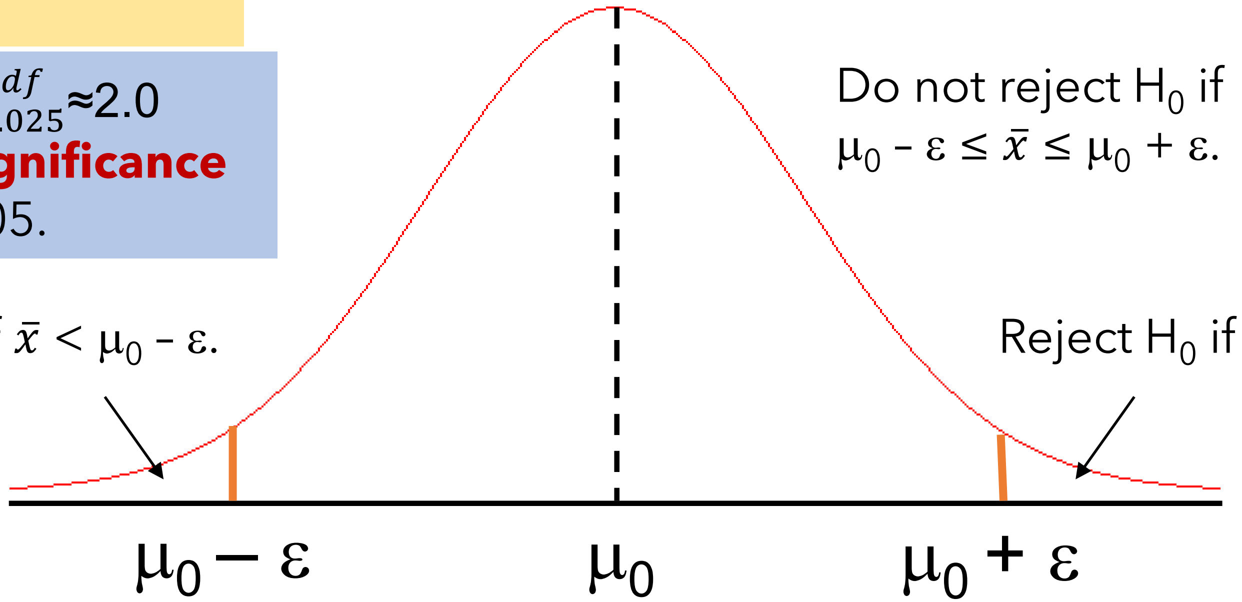
.

Commonly we use $t_{.025}^{df} \approx 2.0$, corresponding to a **significance level** of α = 0.05.

Do not reject H_0 if $\mu_0 - \varepsilon \leq \bar{x} \leq \mu_0 + \varepsilon$.

Reject H_0 if $\bar{x} < \mu_0 - \varepsilon$.

Reject H_0 if $\bar{x} > \mu_0 + \varepsilon$.



Two-Sided Hypothesis Tests

Based on historical data, our team of programmers produces an average of 1000 lines of production-quality code per day. In the last 100 days, our team has used a new integrated development environment, producing a mean of 1070 lines of production-quality code and standard deviation of 300 lines. Can we conclude that the new environment affects programmer productivity?

$H_0 : \mu = 1000$
$H_1 : \mu \neq 1000$

If H_0 was true, $\bar{x}$ would be normally distributed with mean 1000 and standard deviation $300/sqrt(100)$= 30.

The t-score corresponding to $\bar{x}$ = 1070 is t= (1070– 1000) / 30 = 2.33.

Assuming a significance level of α = 0.05 and the corresponding threshold $t_{.05/2}^{df=99} = \pm 1.98$ for a two-sided test, we can reject the null hypothesis and conclude that m ≠ 1000. The new environment does affect productivity!

Suppose α = 0.01 → $t_{.005}^{df=99} = \pm 2.63$; Do we reject?

One-Sided Hypothesis Tests

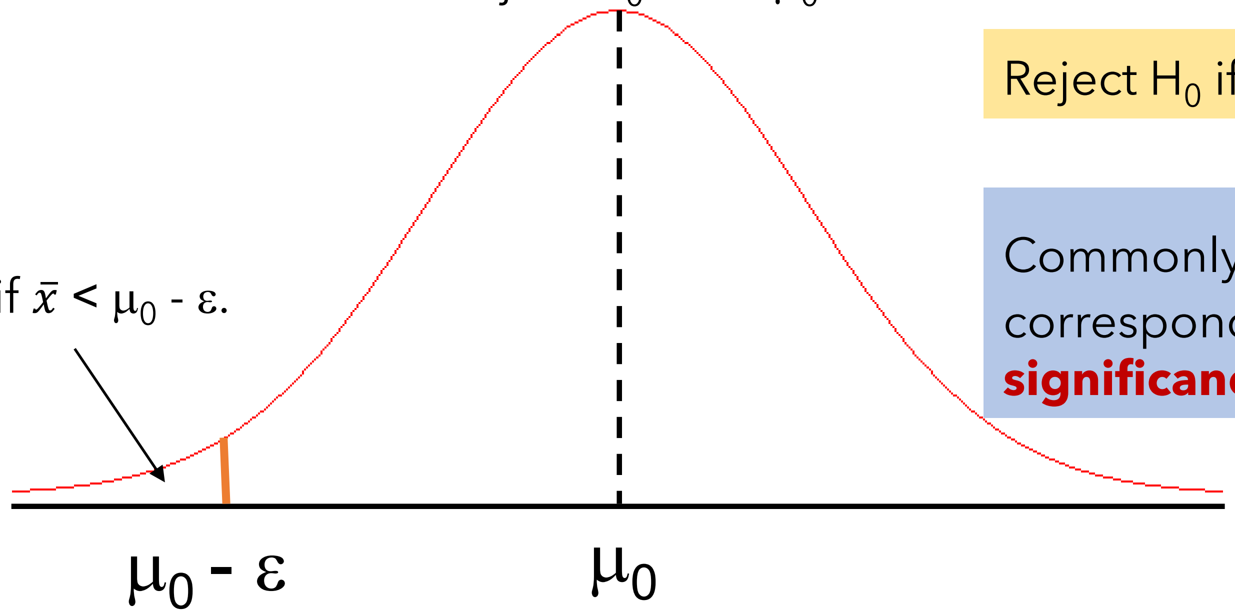
We want to test H_1: $\mu < \mu_0$ against H_0: $\mu \geq \mu_0$.

Solution: reject the null hypothesis if $\bar{x}$ is sufficiently smaller than μ_0.

Assume $\mu = \mu_0$. According to the CLT, $\bar{x}$ is normally distributed with mean μ_0 and standard deviation $\sigma/\sqrt{N} \approx s/\sqrt{N}$. Thus the t-score of $\bar{x}$ is $t = (\bar{x} - \mu_0) / s/\sqrt{N}$

Do not reject H_0 if $\bar{x} \geq \mu_0 - \varepsilon$.

Reject H_0 if $\bar{x} < \mu_0 - \varepsilon$.



Reject H_0 if $t < t_{\alpha}^{df}$

Commonly we use $t_{.05}^{df} \approx -1.65$, corresponding to a **significance level** of $\alpha = 0.05$.

One-Sided Hypothesis Tests

Average days unemployed per year for 36 participants in a job training program was 10 days compared to a national average of 13 days for equivalent untrained individuals. Suppose s = 6. Is this sufficient evidence that the program reduces day unemployed? Test for **α** = 0.05.

$$H_0 : \mu \geq 13$$
$$H_1 : \mu < 13$$

If H_0 was true with **μ** = 13, $\bar{x}$ would be normally distributed with mean 13 and standard deviation $6/\sqrt{36} = 1$.

The z-score corresponding to x = 10 is t = (10 – 13 / 1) = -3.00

We can reject H_0 since $t^{35}_{.05}$ **< -1.69**.

We do have sufficient evidence to conclude that that program participants spend less time unemployed than nonparticipants

Why we conducted the hypothesis test for μ = 13. If we can reject H_0 for μ = 13, we can always reject H_0 for any value of μ >13.

One-Sided Hypothesis Tests

We want to test H_1: $\mu > \mu_0$ against H_0: $\mu \le \mu_0$.

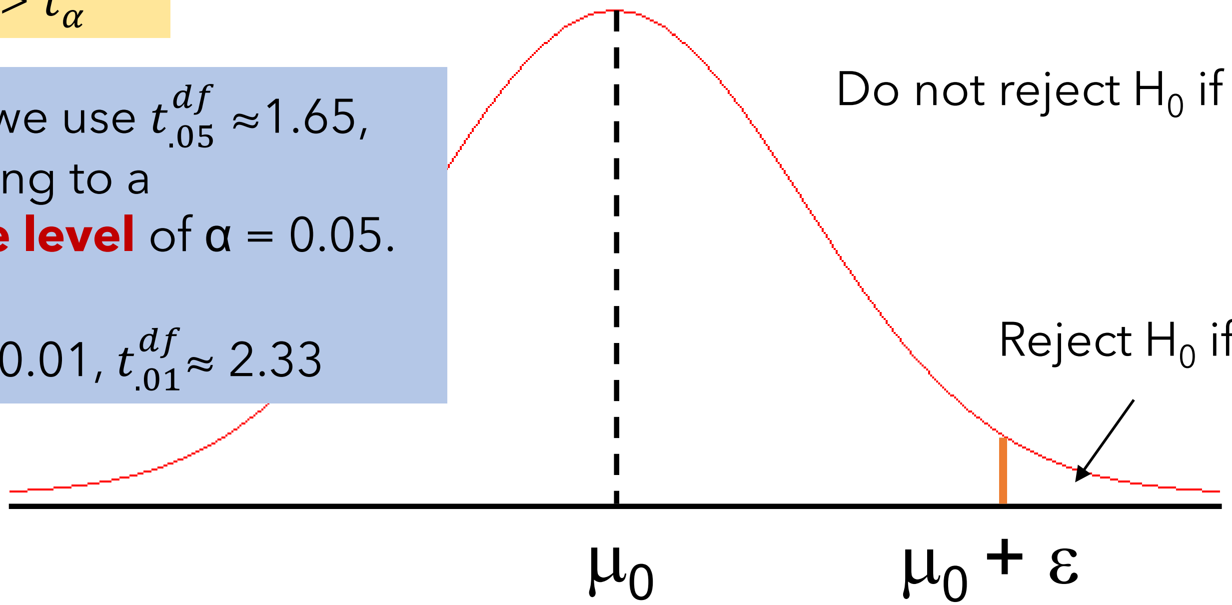
<u>Solution</u>: reject the null hypothesis if $\bar{x}$ is sufficiently larger than μ_0.

Assume $\mu = \mu_0$. According to the CLT, $\bar{x}$ is normally distributed with mean μ_0 and standard deviation $\sigma/\sqrt{N} \approx s/\sqrt{N}$. Thus the t-score of $\bar{x}$ is $t = (\bar{x} - \mu_0) / s/\sqrt{N}$

Reject H_0 if $t > t_\alpha^{df}$

Commonly we use $t_{.05}^{df} \approx 1.65$, corresponding to a **significance level** of $\alpha = 0.05$.

*But for $\alpha = 0.01$, $t_{.01}^{df} \approx 2.33$



One-Sided Hypothesis Tests

A computer supplies retail chain has a policy of only opening stores in communities where households spend more than \$40 per year on computing supplies and equipment. A survey of 100 households in Monroeville finds that average expenditures in the sample are \$40.50 with a standard deviation of \$10. Is this strong evidence that the community spends more than \$40?

$$H_0 : \mu \leq 40$$
$$H_1 : \mu > 40$$

If H_0 was true with $\mu = 40$, $\bar{x}$ would be normally distributed with mean 40 and standard deviation $10/\sqrt{100} = 1$.

The z-score corresponding to $\bar{x} = 40.50$ is t= (40.50 – 40) / 1 = 0.5.

We cannot reject H_0 since **$t \leq 1.66$**.

Assuming a significance level of $\alpha = 0.05$ and the corresponding threshold $t^{99}_{.05} = 1.66$ for a one-sided test, we cannot reject the null hypothesis. We do not have sufficient evidence to conclude that the average household in this community spends more than \$40 each year.

T-statistics versus Z-Score

- Because we always substitute s for σ for hypothesis tests about μ, we should always in principal use t-score cut-off with df=n-1 for these tests.
- T-scores are not used in hypothesis tests about population proportions—the topic of the next slide

Hypothesis Tests About Population Proportions (p)

A survey of 400 customers shows that 43% prefer the new online bill payment system to the old pay-by-mail system. Is this sufficient evidence to show that a minority of customers prefer the new system?

$H_0 : p \geq 0.5$
$H_1 : p < 0.5$

If H_0 was true with p = 0.5, $\hat{p}$ would be normally distributed with mean 0.5 and standard deviation $\sqrt{(p)\,(1-p)/n} = \sqrt{(0.5)\,(0.5)/400} = 0.025$.

Z-scores, not t-scores, are used because there is no substitution of s for σ.
The z-score corresponding to $\hat{p}$ = 0.43 is z = (0.43 – 0.5) / 0.025 = -2.8.

We can reject H_0 since **z < -1.645**.

Assuming a significance level of $\alpha = 0.05$ and the corresponding threshold $z_\alpha = 1.645$ for a one-sided test, we can reject the null hypothesis and conclude that $p < 0.5$. A majority of customers do not prefer the new system!

Test of Understanding

- The marketing manager of a financially strapped theatre company wants to increase the average donation of existing donors. Historically, the average gift of existing donors was $200. To increase this average, he proposes to offer a gift valued at $20 for all such donors who again donate. To test the effectiveness of gift giving, he makes this offer to 144 donors. He finds that their average gift is $230 with a standard deviation of $150. Test the hypothesis that the gift is effective in increasing average giving. Clearly state the null and alternative hypotheses and test for $\alpha=.01$.