Feb 11

Hypothesis Tests & Confidence Intervals Review Activity

Problem 1: State all the steps in a general Hypothesis Test

Step 0 Check req
Step 1 State Hyp
Step 2 LOS (α)
Step 3 Test Statistic
Step 4 P-Val M or Crit Val M
Step 5 Decision & conclusion

PVal Method
P(Critical Region) ← use test statistic (z^* or t^*) as cut-off

Critical Value Method
Critical Region use critical values z_α (or $z_{\alpha/2}$) or t_α (or $t_{\alpha/2}$)
· Make decision based on wheter test statistic is inside/outside critical region

Problem 2: Briefly explain the two methods for making a decision in a Hypothesis Test

PVM use P low, null go (Reject) P high, null fly (FTR)

CVM test statistic is inside (reject) or outside (FTR Ho)

Use the following information to answer problem 3 on the next page

Calculator Functions (unordered!)
- χ^2-GOF-Fit
- 2PropZTest
- TTest
- Anova
- PropZTest
- 2SampTInt
- 2PropZInt
- LinRegTTest
- TInterval
- 2SampTTest

Distribution Used (unordered!)
- F-distribution
- z-distribution (normal)
- χ^2-distribution
- t-distribution (Student's t)

Notation for computing probability
- $\chi^2cdf(a, b, df)$, $df = k - 1$
- $tcdt(a, b, df)$, $df = n - 1$
- $normalcdf(a, b, \mu, \sigma)$
- $Fcdf(F_0, UB, df_{num}, df_{denom})$, $df_{num} = k - 1$, $df_{denom} = k * (n - 1)$

Problem 3: For Each Hypothesis Test state the: (a) Hypotheses; (b) Requirements; (c) Type of distribution used; (d) Test Statistic; (e) Name of Test on Calculator

Types of Hypothesis Tests

One Proportion: population parameter (p)

a. **Hypotheses:**

now, past

$$\begin{cases} H_0: p = p_0 \\ H_A: p < p_0 \text{ or } p \neq p_0 \text{ or } p > p_0 \end{cases}$$

b. **Requirements:**

(1) SRS (2) "5" (3) $np_0 \geq 5$ & $nq_0 \geq 5$ ($q_0 = 1 - p_0$)

c. **Type of distribution:**

z-dist

Prob: normalcdf(a, b, μ, σ)

Crit Val: invNorm(p, μ, σ, Tail)

d. **Test Statistic:**

$$z^* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

all at once

e. **Calculator:**

1PropZTest

One Mean (μ)

a. **Hypotheses:**

now, past

$$\begin{cases} H_0: \mu = \mu_0 \\ H_A: \mu < \mu_0 \text{ or } \mu \neq \mu_0 \text{ or } \mu > \mu_0 \end{cases}$$

b. **Requirements:**

(1) SRS (2) normal dist OR $n > 30$

c. **Type of distribution:**

t-dist

d. **Test Statistic:**

$$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

e. **Calculator:**

TTest (all at once)

for probability: tcdf(a, b, df)

invT(1−α, df) — critical values

(p_1, p_2) ← same time, different groups

Two Proportions $(p_1 - p_2)$

a. **Hypotheses:**

$$\begin{cases} H_0: p_1 = p_2 \\ H_A: p_1 < p_2 \text{ or } p_1 \neq p_2 \text{ or } p_1 > p_2 \end{cases}$$

(both groups are same)

b. **Requirements:**

1. SRS 2. independent 3. $n_1 p_1 \geq 5$, $n_1 q_1 \geq 5$ & $n_2 p_2 \geq 5$, $n_2 q_2 \geq 5$

c. **Type of distribution:**

z-dist

d. **Test Statistic:**

$$z^* = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

e. **Calculator:**

Test: $\boxed{\text{2PropZTest}}$ Prob.: normalcdf Crit Val: invNorm

Two Means (Independent Sample) $(\mu_1 - \mu_2)$

a. **Hypotheses:**

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_A: \mu_1 < \mu_2 \text{ or } \mu_1 \neq \mu_2 \text{ or } \mu_1 > \mu_2 \end{cases}$$

(Better: $\mu_1 - \mu_2 = 0$)

(Better: $\mu_1 - \mu_2 < 0$ $\quad \mu_1 - \mu_2 \neq 0$ $\quad \mu_1 - \mu_2 > 0$)

b. **Requirements:**

1. SRS 2. indep 3. normal or $(n_1 > 30$ & $n_2 > 30)$

c. **Type of distribution:**

t-dist

d. **Test Statistic:**

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

e. **Calculator:**

$\boxed{\text{2SampTTest}}$

Prob: tcdf

Critical Val: invT

Two Means (Dependent Sample)

Matched Pairs

(μ_d)

a. **Hypotheses:**

$$\begin{cases} H_0: \mu_d = 0 \\ H_A: \mu_d < 0 \text{ or } \mu_d \neq 0 \text{ or } \mu_d > 0 \end{cases}$$

b. **Requirements:**

① SRS ② dependent ③ normal or $(n > 30)$

c. **Type of distribution:**

t-dist

d. **Test Statistic:**

$$t^* = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}}$$

e. **Calculator:**

$\boxed{\text{TTest}}$ Prob: tcdf crit val: invT

---

same as one mean

Confidence Intervals

For each of the following: (a) state the point estimate; (b) write the notation for the confidence interval; (c) state the formula for the margin of error; (d) distribution used; and (e) give the calculator command for finding the confidence interval

- **One Proportion**

pt est: $\hat{p}$

CI: $(\hat{p} - E, \hat{p} + E)$

$\boxed{E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}}$

z-dist

$\boxed{\text{1PropZInt}}$

- **One Mean**

pt est: $\bar{x}$

CI: $(\bar{x} - E, \bar{x} + E)$

$\boxed{E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}}$

t-dist

$\boxed{\text{TInterval}}$

- **Two Proportions**

pt est: $\hat{p}_1 - \hat{p}_2 = \hat{\hat{p}}$

CI: $(\hat{\hat{p}} - E, \hat{\hat{p}} + E)$

$\boxed{E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}$

z-dist

$\boxed{\text{2PropZInt}}$

- **Two Means (Independent Sample)**

pt est: $\bar{x}_1 - \bar{x}_2 = \bar{\bar{x}}$

CI: $(\bar{\bar{x}} - E, \bar{\bar{x}} + E)$

$\boxed{E = t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

t-dist

$\boxed{\text{2SampTInt}}$

same

- **Two Means (Dependent Sample)**

pt est: $\bar{d}$

CI: $(\bar{d} - E, \bar{d} + E)$

$\boxed{E = t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}}$

t-dist

$\boxed{\text{TInterval}}$

Mixed Review

Problem 4: For Each problem below:

a. Determine whether you need a Hypothesis Test (HT) or Confidence Interval (CI)
   i. If it is a HT, state both hypotheses
   ii. If it is a CI, state the point estimate
b. Determine which calculator function you need to use
c. Quick practice: use your calculator to either state the conclusion of a hypothesis test or give the confidence interval (no work necessary!)

1. In a random sample of 360 women, 234 favored stricter gun control laws. In a random sample of 220 men, 132 favored stricter gun control laws. Test the claim that the proportion of women favoring stricter gun control laws is higher than the proportion of men favoring stricter gun control laws.

a) • HT vs CI

↳ proportion vs mean

↳ Two: men vs women

women, men: p_1, p_2 $p_1 - p_2$

$H_0: p_1 = p_2$

$H_A: p_1 > p_2$

b) 2PropZTest

(c) $p = 0.113$, $\alpha = 0.05$

$p > \alpha \rightarrow$ $\boxed{\text{Fail to Reject}}$

2. A survey of the mean age of PCC stats students was conducted by randomly surveying one student from each section of Stats 50. The following data were collected: 20, 25, 18, 19, 21, 19, 24, 20, 36, 19, 20, 20, 19, 22. Assume the underlying distribution is approximately normal. Construct a 95% confidence interval for the average age of a stats student (round the values of the interval to 1 decimal place)

3. A survey in the N.Y. Times Almanac finds the mean commute time (one way) is 25.4 minutes for the 15 largest US cities. The Austin, TX chamber of commerce feels that Austin's commute time is less and wants to publicize this fact. The mean for 35 randomly selected Austin, TX commuters is 22.1 minutes with a standard deviation of 5.3 minutes. At the $\alpha = 0.01$ level, test the claim that the Austin, TX commute is less than the mean commute time for the 15 largest US cities.

4. According to an education consultation agency, in 2011 at least 18% of high school students have used a personal tutor. An Introduction to Statistics class in Davies County, KY conducted a hypothesis test at the local high school to determine if the local high school's percentage was higher. One hundred fifty students were chosen at random and surveyed. Of the 150 students surveyed, 32 have used a personal tutor. Reproduce the hypothesis test and state the conclusions.

5. In a sample of 285 randomly selected PCC students it was found that 165 of them have a tattoo. Construct a 97% confidence interval estimate of the percentage of PCC students that have a tattoo (round the values of the interval to three decimal places).

6. In 2017, the total number of cows in California was estimated to be 1.735 million ~~head~~, and milk production per cow per day was estimated at 7.3 gallons. An agricultural department official randomly samples the milk production of 50 cows at various farms throughout California and determines the mean production to be 7.1 gallons per day with a standard deviation of 0.3 gallons. Test the claim that the mean milk production per cow per day today is different from the 2017 level.

a) HT vs CI

prop vs mean

one mean
Two means (Indp)
Two means (Dep)

$\underset{\text{new}}{\mu} = \underset{\text{past}}{\mu_0}$

$H_0: \mu = \overbrace{7.3 \frac{\text{gal}}{\text{cow}}}^{\mu_0}$

$H_A: \mu \neq 7.3$

b) $\boxed{\text{T Test}}$

c) $\boxed{\text{Reject}}$ vs Fail to Reject

$p = 2.05E-5$

$p = 0.0000205$

$\alpha = 0.05$

$p < \alpha \rightarrow$ P low, null go