Lecture 34

Putting Hypothesis Testing into action

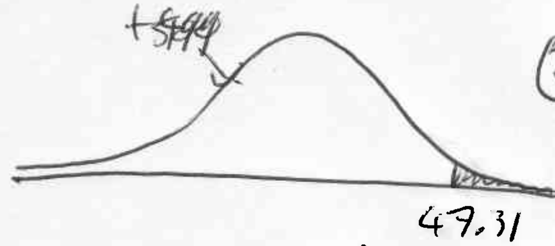
1. A bank wonders whether omitting the annual credit card fee for customers who charge at least \$3000 in a year would increase the amount charged on its credit card. The bank makes this offer to an SRS of 500 of its existing credit card customers. It then compares how much these customers charge this year with the amount that they charged last year. The mean increase is \$565 and standard deviation \$267. Would the bank offer increase the amount charged if they made it to all such customers?

Let μ = mean increase in amount charged if given offer.

$H_0: \mu \leq 0$ (doesn't increase amount charged or even decreases it)

$H_A: \mu > 0$ (increases amount charged).

$$t = \frac{565 - 0}{267/\sqrt{500}} = 47.31$$



$$\text{p-value} = P(T > 47.31) < P(T > 3.373) = .0005$$

Since this p-value is very small we have evidence to reject the null hypothesis. Therefore the offer does increase the amount charged.
Note if $\alpha = .05$ then we would reject if $t > 1.658$
and if $\alpha = .01$ then we would reject if $t > 2.358$

2. A random sample of 9 one-bedroom apartments advertised in a newspaper has these monthly rents (in dollars):

1500, 1650, 1600, 1505, 1450, 1550, 1750, 1200, 2050

Does this data provide evidence that the mean rent of all advertised one-bedroom apartments is higher than \$1600?

Let μ = mean rent of advertised one bedroom apts.

$$\sum x_i = 14255$$

$$\sum x_i^2 = 23007525$$

$$\bar{X} = \frac{14255}{10} = 1425.5$$

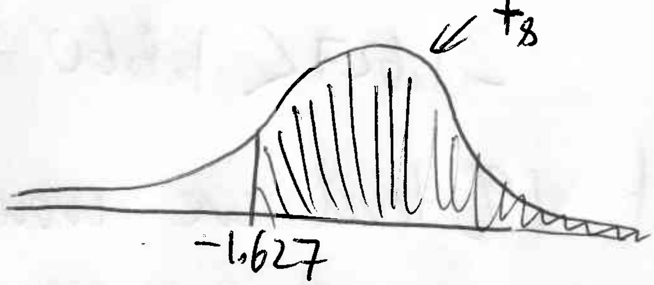
$$s = \sqrt{\frac{23007525 - 10(1425.5)^2}{10-1}} = 231.6217$$

$H_0 : \mu \leq 1600$ (mean rent not above 1600, maybe even less)

$H_A : \mu > 1600$ (above \$1600)

$$t = \frac{1425.5 - 1600}{231.6217/\sqrt{10}} = -1.627$$

$df = 10 - 1 = 9$



$P\text{value} = P(T > -1.627)$ $\geq .5$

$$P(T > -1.627) = P(0 < T < \infty) + P(-1.627 < T < 0)$$

and $P(-1.627 < T < 0) = P(0 < T < 1.627)$ (by symmetry)

$P(0 < T < 1.627) = ?$

using df=8 line from a t-distribution table

$1.397 < 1.627 < 1.86$

ie $.10 > P(T > 1.627) > .05$

so $.40 < P(0 < T < 1.627) < .45$

so $.9 < P(T > -1.627) < .95$

So cannot reject H_0. no evidence to show that mean rent is higher than $1600

Using rejection regions

at 5% level we would not reject since

$-1.627 < 1.860$ (rejection region is $t > 1.860$)

at 1% level we would not reject since

$-1.627 < 2.896$ (rejection region is $t > 2.896$)

3. Statistics can help decide the authorship of literary works. Sonnets by a certain Elizabethan poet are known to contain an average of 6.9 new words (words not used in the poets other known works). The standard deviation in the known works is 2.7. ~~We expect~~ A new manuscript is unearthed with 5 new sonnets which scholars are debating whether or not it is the work of this particular poet. On average the new poems contain ~~an average of~~ 11.2 words not used in the poets other known works. We expect poems by another author to have more new words. Assume that the number of new words is distributed normally. Test whether or not the new sonnets are by this poet.

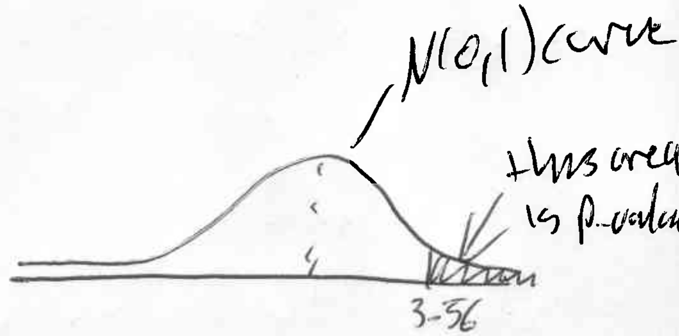
Let μ = number of new words in any newly discovered works of poet.

$\bar{X} = 11.2 \qquad \sigma = 2.7$

$H_0: \mu \leq 6.9$ (probably the authors work)

$H_A: \mu > 6.9$ (probably another authors work).

$$Z = \frac{11.2 - 6.9}{2.7/\sqrt{5}} = 3.56$$



$$\begin{aligned} P\text{value} &= P(Z > 3.56) \\ &\leq P(Z > 3.49) \\ &= 1 - \Phi(3.49) \\ &= 1 - .9998 \\ &= .0002 \end{aligned}$$

Since p-value is small reject the null hypothesis ie it is likely that the sonnets are another authors.

<u>Note</u> rejection regions

at 5% significance level we would reject since
$3.56 > 1.645$ (rejection region is $t > 1.645$)
at 1% significance level we would reject since
$3.56 > 2.326$ (rejection region is $t > 2.326$

⑥

4. The mean yield of corn in the United States is normally about 120 bushells per acre with standard deviation 10. A random sample of 40 farmers this year gives a sample mean of 123.8 bushels per acre. Is this good evidence that the mean yield this year differs from normal?

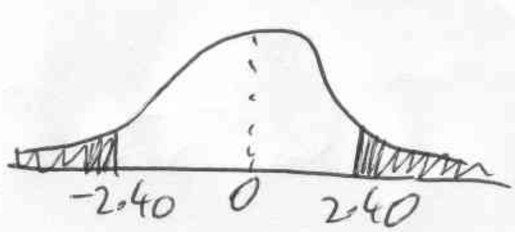
Let μ = mean yield in bushels per acre this year in US.

$\bar{x} = 123.8$

$H_0: \mu = 120$

$H_A: \mu \neq 120$

$$z = \frac{123.8 - 120}{10/\sqrt{40}} = 2.40$$



$$\text{Pvalue} = 2P(Z > |2.40|)$$
$$= 2(1 - \Phi(2.40))$$
$$= 2(1 - .9918) = .0164$$

So we have some evidence to reject the null hypothesis.

Note 1) at 5% level of significance rejection region is

$Z < -1.96$ or $Z > 1.96$ since $2.40 > 1.96$ we would reject H_0 at 5% level of significance.

2) at 1% level of significance rejection region is

$Z < -2.576$ or $Z > 2.576$ since

$-2.576 < 2.40 < 2.576$ we could not reject H_0 at 1% level of significance.

5. The administrator of a nursing home would like to do a time-and-motion study of staff time spent per day performing non-emergency tasks. Historically, 16 person hours per day we spent on these tasks. However she has proposed some new efficiency procedures. and a random set of 12 days after these procedures have been implemented has mean

14.6 hours with standard deviation 2.4 hours. Did the new procedures reduce the time spent on these tasks?

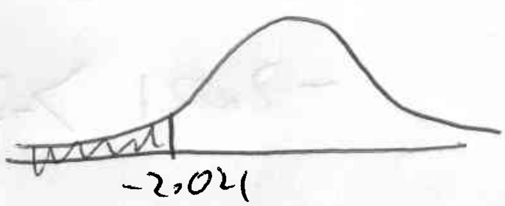
Let μ = mean time spent on non-emergency tasks using new procedures.

$H_0: \mu \geq 16$

$H_A: \mu < 16$

$$t = \frac{14.6 - 16}{2.4/\sqrt{12}} = -2.021$$

$df = 12 - 1 = 11$



P value $= P(T < -2.021)$

$= P(T > 2.021)$ (by symmetry)

From table

$1.796 < 2.021 < 2.201$

$.05 > P(T > 2.021) > .025$

8a

So

$.025 < \text{P-value} < .05$

we have weak evidence to reject the null hypothesis and conclude ~~that~~ time spent on these activities has been reduced using the new procedures.

Note 1) at 5% level of significance we would reject H_0 since

$$-2.021 < -1.796 \quad (\text{rejection region is } t < -1.796)$$

2) at 1% level of significance we could not reject H_0 since

$$-2.021 > -2.718 \quad (\text{rejection region is } t < -2.718)$$

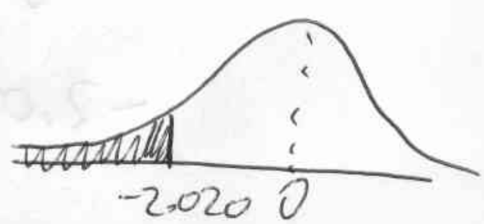
9

6. A buyer wishes to determine whether the mean suger content per orange shipped from a particular grove is less than .027 pounds. A random sample of 50 oranges produceed a mean suger of .025 pounds with standard deviation .007 pounds. Use this data to determine whether or not the mean suger content is less than .027.

$H_0: \mu \geq .027$

$H_A: \mu < .027$

$$t = \frac{.025 - .027}{.007/\sqrt{50}} = -2.020$$



p value $= P(T < -2.020)$

$= P(T > 2.020)$ (by symmetry)

9a

from t table using df=40 line

$1.684 < 2.020 < 2.021$

$.05 > P(T > 2.020) > .025$

ie p-value is between .025 and .05. thus we have weak evidence to reject H_0 in favor of H_A.

Note 1) at 5% level of significance we would reject H_0 since

$-2.020 < -1.684$ (rejection region is $t < -1.684$)

2) at 1% level of significance we would not reject H_0 since

$-2.020 > -2.423$ (rejection region is $t < -2.423$).

10

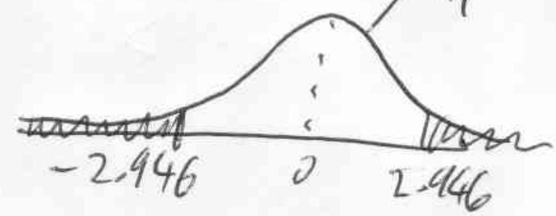
7. A machine produces plastic pipe. Normal pipe has mean diameter 2.5 inches. A random sample of 50 pipes has mean diameter 2.55 inches with standard deviation .12. Do we have evidence that the pipes differ from the standard?

Let μ = mean of pipes comming off machine.

$H_0: \mu = 2.5$ (ie machine producing standard pipes)

$H_A: \mu \neq 2.5$ (ie machine producing pipes out of spec).

$$t = \frac{2.55 - 2.50}{.12/\sqrt{50}} = +2.946$$



$$\text{Pvalue} = 2P(T > |+2.946|)$$
$$= 2P(T > 2.946)$$

10a

From df=40 line

$2.704 < 2.946 < 3.307$

$.005 > P(T > 2.946) > .001$

so $.01 > \text{Pvalue} > .002$

good evidence that we can reject H_0.

ie the pipes are non standard.

Note 1) we would reject at both 5% and 1% levels because

5% reject if $t < -2.021$ or $t > 2.021$

1% reject if $t < -2.704$ or $t > 2.704$.

8. A starting player for an NBA team made only 36.2% of their free throws last season. During the offseason they worked on developing a softer shot in hopes of improving their accuracy. In the first eight games of this season the player made 22 out of 42 free throws. Did they improve their free throw ability?

Let p = proportion of free throws made for this season

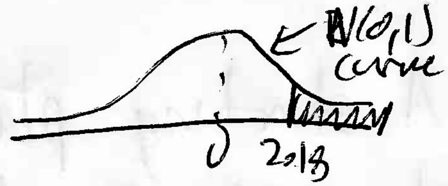
$H_0: p \leq .362$ (no better than last season)

$H_A: p > .362$ (have improved since previous season)

$$\hat{p} = \frac{22}{42} = .524$$

$$z = \frac{.524 - .362}{\sqrt{\frac{.362(1-.362)}{42}}} = 2.184.$$

11a



$$P\text{value} = P(Z > 2.18)$$
$$= 1 - \Phi(2.18)$$
$$= 1 - .9854$$
$$= .0146$$

so we have some evidence to suggest that the player has improved their free throw from last season.

Note 1) at 5% level of significance we would reject H_0 since

$$2.18 > 1.645$$

2) at 1% level of significance we would not reject H_0 since

$$2.18 < 2.326.$$

12

9. A random sample of 160 recent donations at a certain blood bank reveals that 87 were type A blood. Does this suggest that the actual percentage of type A donations differs from 40% (the proportion in the general population.

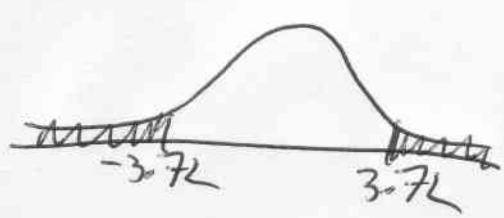
Let p = proportion of type A donations

$$\hat{p} = \frac{87}{160} = .544$$

$H_0: p = .4$

$H_A: p \neq .4$

$$z = \frac{.544 - .4}{\sqrt{\frac{.4(1-.4)}{160}}} = 3.718$$



$$\text{Pvalue} = 2P(Z > |3.72|)$$
$$\leq 2P(Z > 3.49)$$
$$= .0004$$

12d

So we can reject H_0 and therefore a higher proportion of type A people donate blood.