① State: $H_0: p = .2$

$H_a: p > .2$

where p is the true proportion of customers willing to pay for the upgrade. We will use a significance level of $\alpha = .05$

Plan: We will do a 1 sample z test for p.

Random: Random sample of 60 customers

10%: 60 is less than 10% of the companies customers

large counts: $60(.2) = 12 \geq 10$ and $60(.8) = 48 \geq 10$ or 16 success (pay more) and 44 failures (not willing to pay more).

Do:

$$P\left(z > \frac{\frac{16}{60} - .2}{\sqrt{\frac{(.2)(.8)}{60}}}\right) = P(z > 1.29)$$

$\hat{p} = \frac{16}{60} = .27$

using table A, this gives us a pvalue $\approx$ .0985

or using a calculator, 1propztest($p_0 = .2$ (null hypothesis), x= 16 people willing to pay more, n= 60 (sample size), $> p_0$ (alt. hypothesis))

gives us $z = 1.29$ and pvalue $\approx .0984$, $\hat{p} = .27$

Conclude: Since our pvalue of $.0985 > .05$, we fail to reject H_0. We don't have convincing evidence that more than 20% of customers would be willing to pay for the upgrade.

State:

2) H_0: The actual distribution of arrests for spouse abusers like these who will be arrested in 6 months is the same for all 3 police responses.

H_a: The actual distribution of arrests for spouse abusers like these who will be arrested in 6 months is not the same for all 3 police responses.

We will use a significance level of $\alpha = .05$.

Plan: We will use a χ^2 test for homogeneity.

Random: Abusers were randomly assigned to 1 of 3 treatments.

10%: not necessary since this is an experiment and we are not sampling from a population.

Large counts:

$$[B] = \begin{bmatrix} \frac{(543)(212)}{650} = 177 & \frac{(543)(224)}{650} = 187 & \frac{(543)(214)}{650} = 179 \\ \frac{(107)(212)}{650} = 35 & \frac{(107)(224)}{650} = 37 & \frac{(107)(214)}{650} = 35 \end{bmatrix}$$

all expected counts are greater than 5 so large counts is met.

Do: $$\chi^2 = \frac{(187-177)^2}{177} + \frac{(181-187)^2}{187} + \frac{(175-179)^2}{179} + \frac{(25-35)^2}{35} + \frac{(43-37)^2}{37} + \frac{(39-35)^2}{35}$$

$$\approx .56 + .19 + .09 + 2.86 + .97 + .46$$

$$= 5.12$$

$df = (3-1)(2-1) = 2$

5.12

$P(\chi^2 > 5.12)$

using table C with 2 df, gives use a p value between .05 and .01

or on calculator using χ^2-Test (observed = matrix of actual table values $\begin{bmatrix} 187 & 181 & 175 \\ 25 & 43 & 39 \end{bmatrix}$, expected = [B]) gives us $\chi^2 = 5.06$, $df = 2$ and a pvalue $= .0796$.

Conclude: Since our p value is $.0796 > .05\ (\alpha)$ we fail to reject H_0. We don't have convincing evidence the actual distribution of arrests for spouse abusers like these who will be arrested in 6 months is not the same for all 3 police responses.

③ State: $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

$H_a: \mu_1 - \mu_2 < 0$ or $\mu_1 < \mu_2$

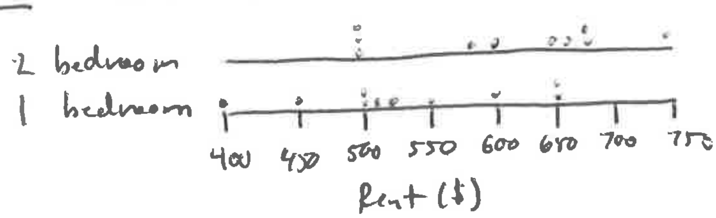
Where μ_1 = true mean rent for 1 bedroom apartments and μ_2 = true mean rent for 2 bedroom apartments on campus. We will use a significance level of $\alpha = .05$.

Plan: We will do a 2 sample t test for $\mu_1 - \mu_2$

Random: 2 independent random samples of 10 apartments

10%: 10 is less than 10% of all 1 bedroom and 2 bedroom apartments near campus.

Large counts: Since n is not less than 30, we will graph to assess skewness & outliers.



There seem to be no outliers and perhaps just a mild skew, so we will assume large counts is met.

Do: using 2 var stat (L_1 = Rent for 1 bedroom, L_2 = rent for 2 bedroom) on my calculator, I set $\bar{x}_1 = \$531$ for 1 bedroom rent and $\bar{x}_2 = \$609$ for 2 bedroom rent and $s_1 = \$82.79$ standard deviation for 1 bedroom rent and $s_2 = \$89.31$ for 2 bedroom rent standard deviation

$$P\left(t < \frac{(531-609)-0}{\sqrt{\frac{82.79^2}{10} + \frac{89.31^2}{10}}}\right) = -2.03 \qquad df = 10-1 = 9$$

$$P(t < -2.03)$$

this gives a p value between .025 and .05

or

on my calculator, using 2 Samp T test (L_1 = rent for 1 bedroom, L_2 = rent for 2 bedroom, $< \mu_2$ (alt hypothesis)) gives us $t = -2.03$, $df = 17.89$, pvalue $\approx .029$.

Conclude: Since our pvalues of $.029 < .05$ (α) we reject H_0. We have convincing evidence the true mean rent of 1 bedroom apartments is less than the true mean rent of 2 bedroom apartments.

(4) <u>State</u>: H_0: The distribution of gas types is the same as claimed by the distributor

H_a: the distribution of gas types is not the same as claimed by the distributor.

We will use a significance level of $\alpha = .05$

<u>Plan</u>: We will do a χ^2 GOF test

<u>Random</u>: Random Sample of drivers

<u>10%</u>: 400 is less than 10% of all customers at the gas station.

<u>Large Counts</u>: Expected counts:

Regular	Premium	Supreme
$(400)(.6)$	$(400)(.2)$	$(400)(.2)$
$=240$	$=80$	$=80$

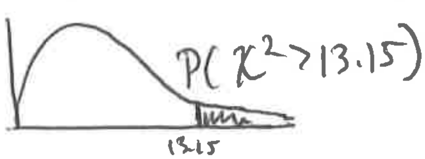
all expected counts are more than 5.

Do: $\chi^2 = \frac{(261-240)^2}{240} + \frac{(51-80)^2}{80} + \frac{(88-80)^2}{80}$

$\approx 1.84 + 10.51 + .8$

$= 13.15$

$df = 3-1 = 2$ This give us a pvalue between .001 and .0025.



or

On my calculator using χ^2 GOF-Test ($L_1 = \{261, 51, 88\}$ observed counts, L_2 = Expected counts, df = 2)

This gives us $\chi^2 = 13.15$, and pvalue $\approx .0014$

<u>Conclude</u>: Since our pvalue of $.0014 < .05$ (α) we reject H_0. We have convincing evidence the distribution of gas types is not the same as claimed by the distributor.

(5) State: $H_0: p_1 - p_2 = 0$ or $p_1 = p_2$

$H_a: p_1 - p_2 > 0$ or $p_1 > p_2$

where p_1 = true proportion of cars that have the brake defect in last year's model.

p_2 = true proportion of cars that have the brake defat in this year's model.

we will use a significance level of $\alpha = .05$.

Plan: we will do a 2 sample z test for $p_1 - p_2$.

Random: Independent random samples, 100 last years cars and 350 of this years cars.

10%: 100 is less than 10% of all of last year's cars

350 is less than 10% of all of this years cars

large counts: 20 of last years cars had a defect (success) and 80 had no defect (failure)

50 of this years cars had a defect (success) and 300 had no defect (failure)

All successes and failures are larger than 10.

Do: $\hat{p}_1 = \frac{20}{100} = .2$ $\quad \hat{p}_2 = \frac{50}{350} = .14$ $\quad \hat{p}_c = \frac{20+50}{100+350} = .16$

$$P\left(z > \frac{(.2-.14)-0}{\sqrt{\frac{(.16)(.84)}{100} + \frac{(.16)(.84)}{350}}}\right) = P(z > 1.44)$$

This gives us a pvalue of $1 - .9251 = .0749$

or on calculator, using 2 Prop Z Test ($x_1 = 20$ defects on last years cars, $n_1 = 100$ sample size of cars last year, $x_2 = 50$ defects on this years cars, $n_2 = 350$ sample size cars this year, $> p_2$ alt. hyp.)

gives us $z = 1.39$ and a pvalue $= .082$.

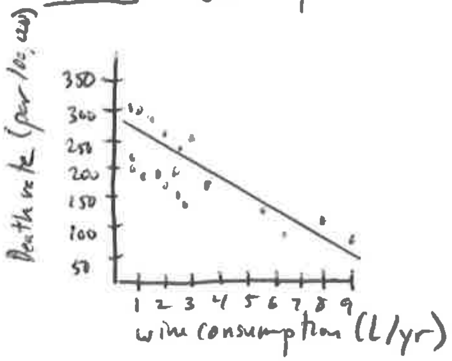
Conclude: Since our p value of $.082 > .05$ (α) we fail to reject H_0. We don't have convincing evidence the true proportion of brake defects is greater in last year's models than this years models.

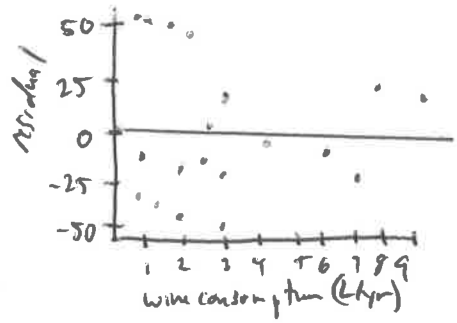
⑥ <u>State</u>: $H_0: \beta = 0$
$H_a: \beta < 0$

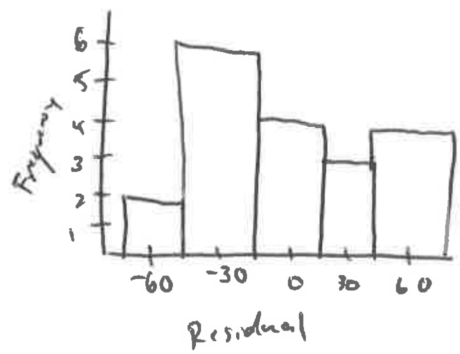
where β = true slope of the population regression line relating heart disease death rate and wine consumption. We will use a significance level of $\alpha = .05$.

<u>Plan</u>: We will do a t test for β

<u>Linear</u>: Scatterplot is roughly linear and the residual plot is well scattered.







<u>Independent</u>: 19 is less than 10% of all countries

<u>Normal</u>: The histogram of residuals has no outlier nor strong skewness

<u>Equal SD</u>: The residual plot shows that smaller wine consumption may have slightly larger residuals ~~than~~ than higher wine consumption, but not a large enough difference to be concerning.

<u>Random</u>: Random sample of 19 countries

<u>Do</u>: using linreg a+bx on calculator, $b = -22.97$.
on calculator, using LinReg TTest (Xlist = wine consumption (L/yr), Ylist = death rate per 100,000, <0 alt. hypothesis) &

This gives us $t = -6.46$, $df = 17$, and pvalue $= .000003$.

<u>Conclude</u>: Since our pvalue of $.000003 < .05\ (\alpha)$ we reject the H_0.
We have convincing evidence of a negative linear relationship between wine consumption and heart disease death rate.

(7) <u>State</u>: $H_0: \mu = \$158$ where μ = true mean amount spent on food by households in this city. We will use a significance level of $\alpha = .05$.

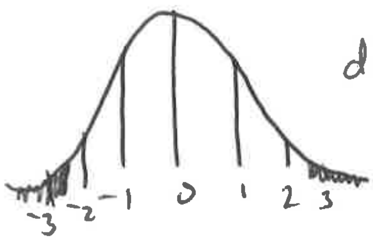
$H_a: \mu \neq \$158$

<u>Plan</u>: <u>Random</u>: random sample of 50 households

<u>10%</u>: 50 is less than 10% of all households in this city

<u>Large Counts</u>: $n = 50 > 30$.

<u>Do</u>: $P\left(t = \frac{165-158}{\frac{20}{\sqrt{50}}}\right) \approx P(t = 2.47)$



$df = 50 - 1 = 49$

using table B, this gives us, using $df = 40$, a pvalue between .01 and .02

or

on calculator using TTest ($\mu_0 = 158$ null hypothesis, $\bar{x} = \$165$ sample mean spent, $s_x = \$20$ sample standard deviation of amount spent, $n = 50$ sample size, $\neq \mu_0$ alternative hypothesis)

This gives us $t = 2.47$, $df = 49$, and pvalue = .0168.

<u>Conclude</u>: Since our pvalue of $.0168 < .05$, we reject H_0. We have convincing evidence that the true mean amount spent on food by households in this city is not \$158.