SUN

10

11 Review 1, 2, 4-6

1. a) independent

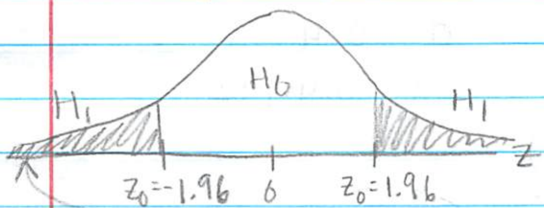
b) normal z, $n_1 > 30$, $n_2 > 30$

a) $\alpha = 0.05$

$H_0: \overset{30\,sec}{\mu_1} = \overset{60\,sec}{\mu_2}$ There is no difference in the avg ratings for 30 second and 60 second ads.

$H_1: \overset{30\,sec}{\mu_1} \neq \overset{60\,sec}{\mu_2}$

Conditions
- SRS
- independent
- $n < \frac{1}{10} N$



$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{(6.03 - 7.09) - 0}{\sqrt{\frac{1.2^2}{142} + \frac{1.3^2}{47}}}$$

$$= \boxed{-4.94}$$

We fail to reject H_1 and reject H_0, $\alpha = 0.05$

$\exists x$ sufficient statistical evidence to suggest that that there is a difference in the avg ratings for 30 sec and 60 sec ads.

P value: $7.94 \times 10^{-7} \approx 0 < \alpha (0.05)$. strong evidence against null.

b) 95%

$$E = z_c \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= 1.96 \sqrt{\frac{1.2^2}{142} + \frac{1.3^2}{47}}$$

$$= 0.4208$$

$$-1.06 - 0.4208 < \mu_1 - \mu_2 < -1.06 + 0.4208$$

$$-1.4808 < \mu_1 - \mu_2 < -0.6392$$

If we repeated this 1000 times with the same sample sizes ($n_1 = 142$, $n_2 = 47$), we expect to capture the difference between μ_1 and μ_2 950 times. For this particular sample, we got an interval of -1.4808 to -0.6392. Since all the values are negative, the avg ratings of 30 sec ads are lower than those of 60 sec ads.

2. dependent

t distribution

$H_0 = \mu_d = 0$ there is no difference in the rat's time needed to complete the maze between smaller and larger rewards.

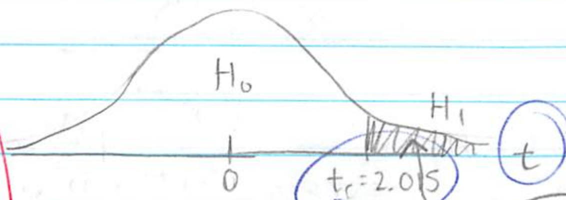
Conditions
- SRS
- dependent
- $n <$ to N

use t's

$H_1 : \mu_d > 0$

$\bar{d} = 0.4$

$s_d = 0.4472$



$d \to t = \dfrac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \dfrac{0.4 - 0}{\frac{0.4472}{\sqrt{6}}} = \boxed{2.19}$

we fail to reject H_1 and reject H_0, $\alpha = 0.05$

∃ X sufficient statistical evidence to suggest that rats receiving larger rewards tend to run the maze in less time.

p value = 0.04 < 0.05 (α), strong evidence against null

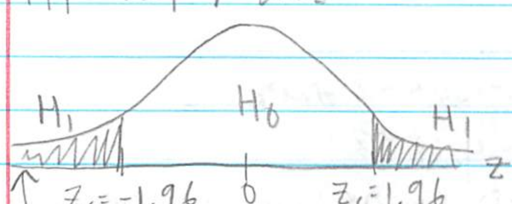
4. independent

normal z distribution

$H_0 : \underset{\text{BUS A}}{\mu_1} = \underset{\text{BUS B}}{\mu_2}$ there is no difference in off-schedule times for Bus line A and B from Denver to Durango.

Conditions
- SRS
- independent
- $n <$ to N

$H_1 : \mu_1 \neq \mu_2$



$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

$= \dfrac{(53 - 62) - 0}{\sqrt{\frac{19^2}{81} + \frac{15^2}{100}}}$

$= \boxed{-3.4752}$

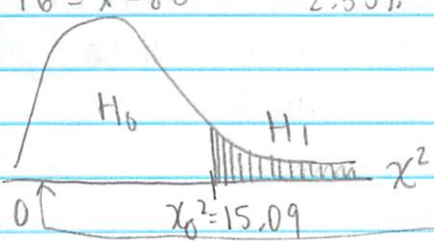
p value ≈ 0 < α, strong evidence against the null

we fail to reject H_1 and reject H_0, $\alpha = 0.05$

∃ X sufficient statistical evidence to suggest there is a significant difference in off-schedule times of Bus line A and B from Denver to Durango.

		O	E	$(O-E)^2$	$(O-E)^2/E$
$56 \le x < 60$	2.35%	14	$0.0235 \times 620 = 14.57$	0.3249	0.0223
$60 \le x < 64$	13.5%	86	$0.135 \times 620 = 83.7$	5.29	0.0632
$64 \le x < 68$	34%	207	$0.34 \times 620 = 210.8$	14.44	0.0685
$68 \le x < 72$	34%	215	210.8	17.64	0.0837
$72 \le x < 76$	13.5%	83	83.7	0.49	0.0059
$76 \le x < 80$	2.35%	15	14.57	0.1849	0.0127

$\chi^2 = \Sigma = \boxed{0.2563}$



df = E-1 = 6-1 = 5

$\alpha = 0.01$

$\chi_0^2 = 15.09$

we fail to reject H_0, we reject H_1, $\alpha = 0.01$

∃x insufficient statistical evidence to suggest that the population of avg daily Jan temps in Hana, Maui does not follow a normal distribution. Thus, it does follow the pattern.

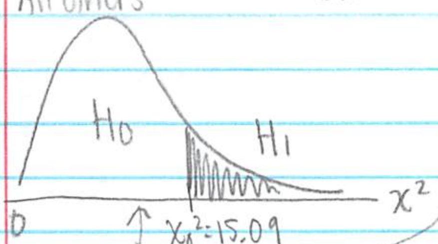
P value is greater than $0.995 > \alpha\ (0.01)$, weak evidence against null

9. $H_0: \chi^2 = 0$ the population of the census distribution follows the sample distribution.

$H_1: \chi^2 > 0$

		O	E	$(O-E)^2$	$(O-E)^2/E$
Black	10%	127	$0.1 \times 1215 = 121.5$	30.25	0.249
Asian	3%	40	$0.03 \times 1215 = 36.45$	12.60	0.3457
Anglo	38%	480	$0.38 \times 1215 = 461.7$	334.89	0.7253
Latino/a	41%	502	$0.41 \times 1215 = 498.15$	14.82	0.0298
Native American	6%	56	$0.06 \times 1215 = 72.9$	285.61	3.92
All others	2%	10	$0.02 \times 1215 = 24.3$	204.49	8.42

$\chi^2 = \boxed{\Sigma = 13.69}$



df = E-1 = 6-1 = 5

$\alpha = 0.01$

we fail to reject H_0 and reject H_1, $\alpha = 0.01$

∃x insufficient statistical evidence to suggest that the census distribution varies from the sample distribution. Thus, the pop fits the distribution.

P value is between 0.01 and $0.025 > \alpha\ (0.01)$, weak evidence against null of customer ages in snoop report

10. $H_0: \chi^2 = 0$ The pop fits the distribution of the sample report.

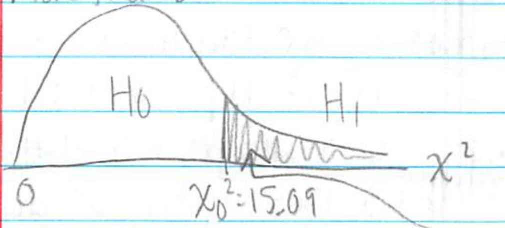
$H_1: \chi^2 > 0$

		O	E	$(O-E)^2$	$(O-E)^2/E$
Less than 14	12%	88	$0.12 \times 519 = 62.28$	661.52	10.62
14-18	29%	135	$0.29 \times 519 = 150.51$	240.56	1.60
19-23	11%	52	$0.11 \times 519 = 57.09$	25.91	0.4538
24-28	10%	40	$0.10 \times 519 = 51.9$	141.61	2.73
29-33	14%	76	$0.14 \times 519 - 72.66$	11.15	0.1536
More than 33	24%	128	$0.24 \times 519 = 124.56$	3.44	0.0276

$df = E - 1 = 6 - 1 = 5$ $\chi^2 = \Sigma = \boxed{15.585}$

$\alpha = 0.01$

$\chi_0^2 = 15.09$



we fail to reject H_1 and reject H_1, $\alpha = 0.01$

∃x sufficient statistical evidence to suggest that the distribution of customer ages difers with the sample report. The pop does not fit the distribution.

P value is between 0.005 and $0.01 < \alpha(0.01)$, strong evidence against null

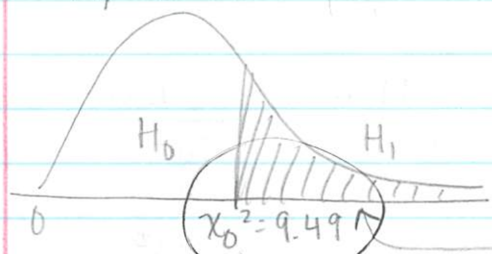
SUN

10

12.2 (2, 4-6, 9, 10)

2. $H_0: \chi^2 = 0$ population of U.S. households fit the distribution

$H_1: \chi^2 > 0$



d.f. = E-1 = 5-1 = 4

$\alpha = 0.05$

$\chi_0^2 = 9.49$

		O	E	$(O-E)^2$	$(O-E)^2/E$
Married w/ kids	26%	102	0.26×411 = 106.86	23.62	0.2210
Married w/o kids	29%	112	0.29×411 = 119.19	51.7	0.4338
single parent	9%	33	0.09×411 = 36.99	15.92	0.4304
one person	25%	96	0.25×411 = 102.75	45.56	0.4434
other	11%	68	0.11×411 = 45.21	519.38	11.49

$\chi^2 = \Sigma = \boxed{13.02}$

we fail to reject H_1 and reject H_0, $\alpha = 0.05$

∃X significant statistical evidence to suggest that the distribution of U.S. households does not fit the Dove creek distribution.

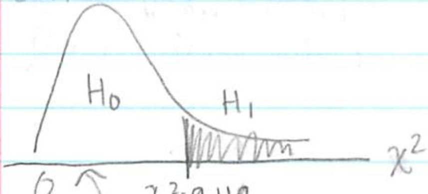
P value is between 0.01 and 0.025 < α (0.05) strong evidence against the null

4. $H_0: \chi^2 = 0$ the population of type of browse favored fits the deer feeding distribution.

$H_1: \chi^2 > 0$

		O	E	$(O-E)^2$	$(O-E)^2/E$
sage brush	32%	102	0.32×320 = 102.4	0.16	0.0016
Rabbit brush	38.7%	125	0.387×320 = 123.84	1.35	0.0109
salt brush	12%	43	0.12×320 = 38.4	21.16	0.5510
service berry	9.3%	27	0.093×320 = 29.76	7.62	0.2560
other	8%	23	0.08×320 = 25.6	6.76	0.2641

$\chi^2 = \Sigma = \boxed{1.0836}$



df = E-1 = 5-1 = 4

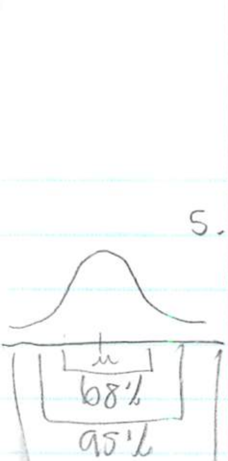
$\alpha = 0.05$

$\chi_0^2 = 9.49$

we fail to reject H_0 and reject H_1, $\alpha = 0.05$

∃X insufficient statistical evidence to suggest that the distribution of type of browse favored doesn't fit the deer feeding distribution. Thus, it fits the distribution

P value is between 0.1 and 0.9 > α (0.05) weak evidence against null



in Kitcarson, Colorado

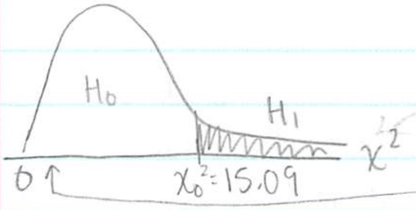
5. i) column 1 are different ranges of temperatures, but expressed using μ and σ. when converting column 1 to 2, plug in 75 for μ and 8 for σ to find the range in Fahrenheit. column 3 is the expected percentage of time that x falls in that specific region. It is the normal distribution

ii) $H_0: \chi^2 = 0$ the population of average daily July temperature fits a normal distribution.

$H_1: \chi^2 > 0$

		O	E	$(O-E)^2$	$(O-E)^2/E$
$51 \le x < 59$	2.35%	16	$0.0235 \times 620 = 14.57$	2.04	0.1400
$59 \le x < 67$	13.5%	78	$0.135 \times 620 = 83.7$	32.49	0.3882
$67 \le x < 75$	34%	212	$0.34 \times 620 = 210.8$	7.84	0.0372
$75 \le x < 83$	34%	221	210.8	104.04	0.4935
$83 \le x < 91$	13.5%	81	83.7	7.29	0.0871
$91 \le x < 99$	2.35%	12	14.57	6.60	0.4530

$\chi^2 = \Sigma = \boxed{1.57}$



$df = E - 1 = 6 - 1 = 5$

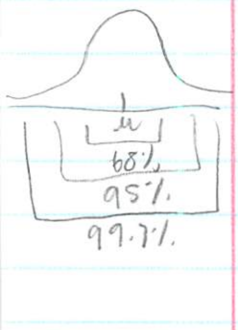
$\alpha = 0.01$

$\chi_0^2 = 15.09$

we fail to reject H_0 and reject H_1, $\alpha = 0.01$

$\exists$ x insufficient statistical evidence to suggest that the pop of avg daily July temp in Kitcarson, Colorado does not fit a normal distribution. In other words, it fits the distribution.

P value is between 0.9 and 0.95 $> \alpha (0.01)$, weak evidence against null



6. i) column 1 refers to all the different regions under the normal curve, represented by x's and σ's. It is a temperature range, and so is column 2 but it has plugged in $\mu = 68$ and $\sigma = 4$ to obtain numerical temperature ranges in Fahrenheit. column 3 is the expected % of time to get results in a specified temp region. It is the normal distribution.

ii) $H_0: \chi^2 = 0$ The pop of avg daily Jan temps in Hana, Maui follows a normal distribution.

$H_1: \chi^2 > 0$

9. a) $H_0: \chi^2 = 0$ Ticket sales and type of billing are independent.

$H_1: \chi^2 > 0$

b) 7.52

c) $\chi^2 = 1.868$

d) P value = 0.600 > 0.05 (α), thus we reject to fail the null hypothesis

2.

	T		F	
C	114	238.39	420	295.61
MD	785	715.63	818	887.375
Law	176	120.98	95	150.02

cell	O	E	O-E	$(O-E)^2$	$(O-E)^2/E$
1	114	238	-124	15376	64.6
2	420	295	125	15625	52.97
3	785	715	70	4900	6.85
4	818	887	-69	4761	5.37
5	176	120	56		17.82
6	95	150	-65		20.17

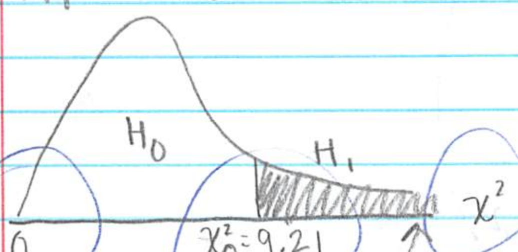
$\chi^2 = \Sigma (O-E)^2/E$

$= \boxed{167.72}$

174

$H_0: \chi^2 = 0$ occupations and personality preferences are independent.

$H_1: \chi^2 > 0$



d.f. = (R-1)(C-1) = 2-1 = 2

$\alpha = 0.01$

$\chi^2_0 = 9.21$

We fail to reject H_1 and reject H_0, $\alpha = 0.01$

∃x sufficient statistical evidence to suggest that occupations and personality preferences are not independent.

P value < 0.005 ≈ 0

7. $H_0: \chi^2 = 0$ Age and type of movie preferred are independent.

$H_1: \chi^2 > 0$

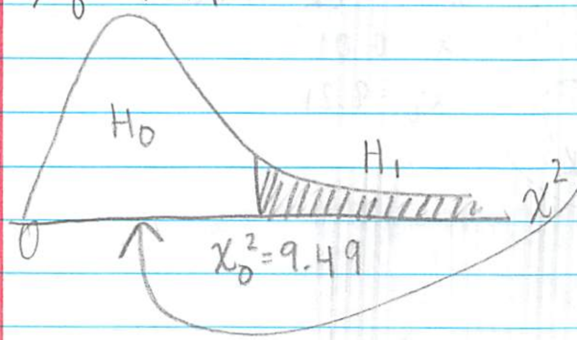
cell	O	E	O−E	$(O-E)^2$	$(O-E)^2/E$
1	8	10.60	−2.60	6.76	0.64
2	15	12.06	2.94	8.64	0.72
3	11	11.33	−0.33	0.11	0.0097
4	12	9.35	2.65	7.02	0.75
5	10	10.65	−0.65	0.42	0.039
6	8	10	−2	4	0.4
7	9	9.04	−0.04	0.0016	≈0
8	8	10.29	−2.29	5.24	0.51
9	12	9.67	2.33	5.43	0.56

$\chi^2 = \Sigma \frac{(O-E)^2}{E} = \boxed{3.62}$

d.f. $= (R-1)(C-1) = 2 \cdot 2 = 4$

$\alpha = 0.05$

$\chi^2_0 = 9.49$



Fail to reject H_0, reject H_1, $\alpha = 0.05$

∃ insufficient statistical evidence to suggest that the age and type of movie preferred are not independent.

Thus, they are independent.

P value is between 0.10 and 0.90 > α (0.05)

weak evidence against null

SUN

Chi-Squared SG

1. Read p514-523

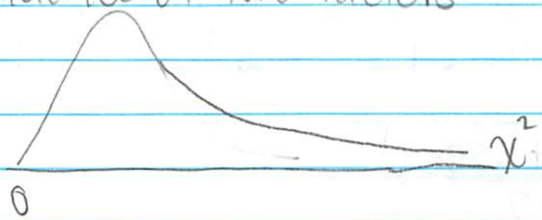
2\. 4x6 contingency table

3\.

*					

4\. independence of two factors

5\. False



6\. False

7\. $n-2$

8\. Go to table six, find designated d.f., and find critical value for specified area α in the right tail of the distribution

9\. $E = \frac{(\text{Row total})(\text{column total})}{\text{sample size}}$

10\. $\frac{(100)(60)}{300} = 20$

11\. 0

12\. χ^2 measures the sum of differences between observed frequency O and expected frequency E

13\. $\chi^2 = \Sigma \frac{(O-E)^2}{E}$

$= 0.04 + 2.50 + 5.06 + 1.00 + 2.02 + 1.04 + 0.83$

$= 13.31$

14\. No, still need to test for independence

15\. $d.f. = (R-1)(C-1)$

16\. $d.f. = (4-1)(6-1)$

$= 15$

SUN

10

$\chi^2 = 0$ 12-1 (1, 2, 7, 9)

1. H_0: Occupations and personality preferences are independent.
$\chi^2 > 0$
H_1: Occupation and personality preferences are not independent

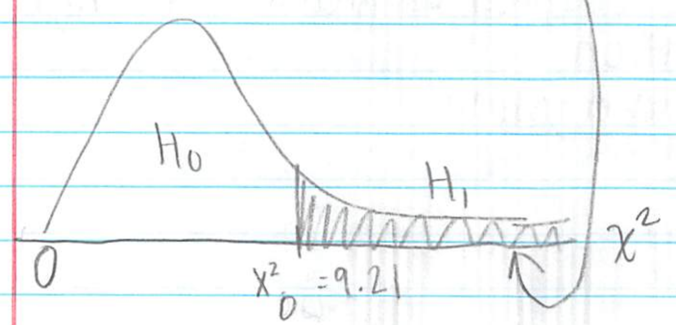
cell	O	E	O − E	$(O-E)^2$	$(O-E)^2/E$
1	308	241.05	66.95	4482.30	18.59
2	226	292.95	−66.95	4482.30	15.30
3	667	723.61	−56.61	3204.69	4.43
4	936	879.39	56.61	3204.69	3.64
5	112	122.33	−10.33	106.71	0.8723
6	159	148.67	10.33	106.71	0.7178

$\chi^2 = \Sigma \frac{(O-E)^2}{E} = \boxed{43.5501}$

d.f. $= (2)(1) = 2$

$a = 0.01$

$\chi^2_0 = 9.21$



Fail to reject H_1, reject H_0, $a = 0.01$
∃X sufficient statistical evidence to suggest that the Meyers-Briggs personality preference and occupations are not independent.

P value < 0.005 ≈ 0 strong evidence against null

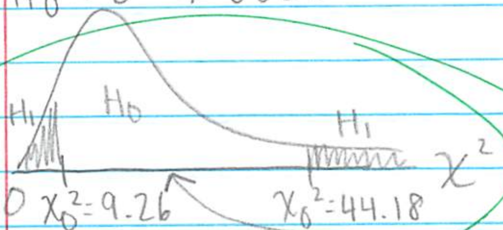
8. $n = 24$ $df = 23$ "different" $\alpha = 0.01$

$\sigma = 75$ $\sigma^2 = 5625$

$s = 72$ $s^2 = 5184$

$H_0: \sigma^2 = 5625$ The variance of a Midwest state's attorney scores on the preliminary bar exam is the same.

$H_0: \sigma^2 \neq 5625$



$1 - 0.005 = 0.995$ $\chi^2 = 9.26$

$0.005: \chi_0^2 = 44.18$

$\chi^2 = (n-1)s^2/\sigma^2$

$= 23 \cdot 5184 / 5625$

$= \boxed{21.20}$

we fail to reject H_1 and reject H_1, $\alpha = 0.01$

$\exists x$ insufficient statistical evidence to suggest that the variance of scores of a midwest state attorneys on the preliminary bar exam is different. Thus, the variance is the same.

P value is between 0.1 and 0.9 < 0.01 (α), weak evidence against the null

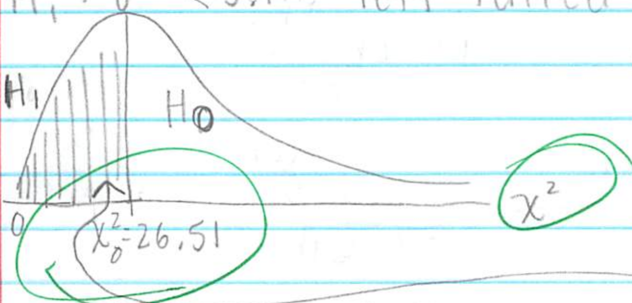
$p = 0.637$

cathy sun
4/17/19

10

12.3 (2, 5-8)

2. $n=41$ $df=40$ $\alpha=0.05$
$s^2=3.3$ $\sigma^2=5.1$ "less than"
$H_0: \sigma^2=5.1$ the variance in age in years of a rural quebec woman at the time of her first marriage is the same.
$H_1: \sigma^2<5.1$ left-tailed test

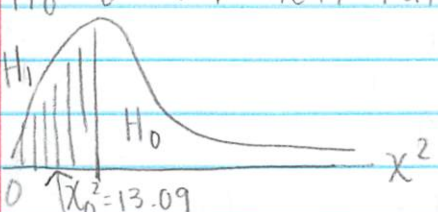


$\alpha=0.05$ $1-0.05=0.95$
$\chi_0^2=26.51$

$\chi^2=(n-1)s^2/\sigma^2$
$=40\times 3.3/5.1$
$=\boxed{25.88}$

we reject H_0 and fail to reject H_1, $\alpha=0.05$.
∃X sufficient statistical evidence to suggest that the variance in age in years of a rural quebec woman at the time of her first marriage is less than the current variance of 5.1.
P value is between 0.025 and 0.05 $<\alpha(0.05)$, strong evidence against null

5. $n=24$ $df=23$ $\alpha=0.05$
$s=1.9$ $s^2=3.61$ "smaller"
$\sigma=3$ $\sigma^2=9$
$H_0: \sigma^2=9$ the variance in the new typhoid shot protection times is the same.
$H_0: \sigma^2<9$ left-tailed test



$\alpha=0.05$ $1-0.05=0.95$
$\chi_0^2=13.09$
$\chi^2=(n-1)s^2/\sigma^2$
$=23(3.61)/9$
$=\boxed{9.23}$

P value is <0.005, $<\alpha(0.05)$, strong evidence against null

we reject H_0 and fail to reject H_1, $\alpha=0.05$
∃X sufficient statistical evidence to suggest that the variance in the new typhoid shot protection times is smaller.

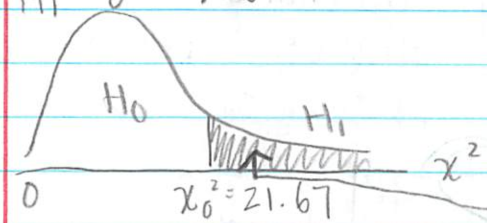
6. $n=10$ $df=9$ $\alpha=0.01$ "larger"

$s=24$ $s^2=576$

$\sigma=15$ $\sigma^2=225$

$H_0: \sigma^2=225$ The variance in prize bull's return to normal is the same.

$H_1: \sigma^2 > 225$



$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{9(576)}{225} = \boxed{23.04}$$

we reject H_0 and fail to reject H_1, $\alpha=0.01$

∃x sufficient statistical evidence to suggest that prize bull's return to normal time in vermont is larger than that of the journal.

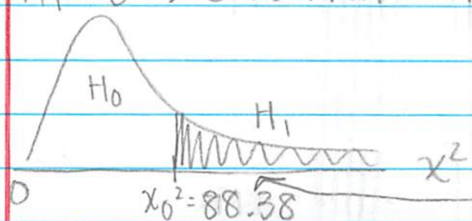
p value is between 0.005 and 0.01 > α (0.01), strong evidence against the null

7. $n=61$ $n=60$ $\alpha=0.01$

$s^2=0.27$ $\sigma^2=0.15$ "all" ∴ all exceed ∴ right-tailed test

$H_0: \sigma^2=0.15\ mm^2$ The variance in all the engine fan blades on commercial jets is the same.

$H_1: \sigma^2 > 0.15\ mm^2$ right-tailed test



$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{60 \times 0.27}{0.15} = \boxed{108}$$

we reject H_0 and fail to reject H_1, $\alpha=0.01$

∃x sufficient statistical evidence to suggest that the variance in all the engine fan blades on the commercial jet is larger than 0.15 mm^2 and all blades must be replaced

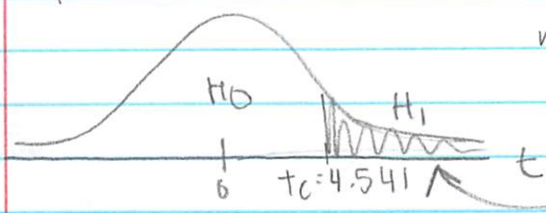
P value is less than 0.005 < 0.01 (α), strong evidence against the null

c) $H_0: \beta = 0$ The population slope is 0.

$H_1: \beta > 0$

$n = 5$ $df = 3$ $\alpha = 0.01$

$t = \boxed{7.93}$



we reject H_0 and fail to reject H_1, $\alpha = 0.01$

∃x sufficient statistical evidence to suggest that there is a positive slope between the list price for a random selection of cars of different models/options and the dealer invoice.

p value = 0.0021 < 0.01 (α), strong evidence against null

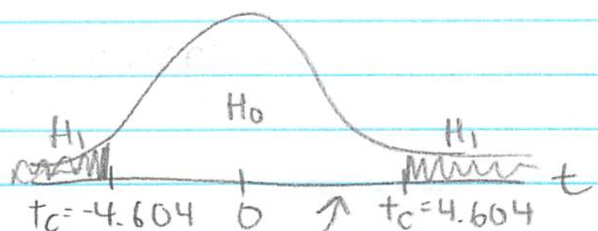
7a) $n = 6$ $df = 4$ $\alpha = 0.01$ two-tailed

$t_c = 4.604$

$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

$= \frac{0.9\sqrt{6-2}}{\sqrt{1-0.9^2}}$

$= \boxed{4.1295}$



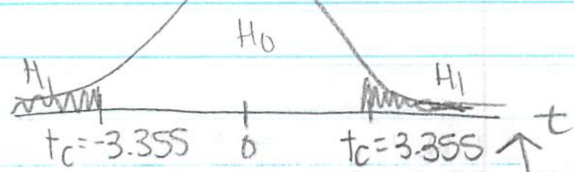
r is not statistically significant

b) $n = 10$ $df = 8$ $\alpha = 0.01$ two-tailed

$t_c = 4.604$

$t_c = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.9\sqrt{10-2}}{\sqrt{1-0.9^2}}$

$= \boxed{5.84}$



r is statistically significant

c) They are different because of sample size, which does play an important part. As n increases, the critical t score needed decreases. On the other hand, the sample statistic t value increases as n increases. Thus, results become more and more statistically significant as sample size increases. In addition, p decreases, providing further significant evidence.

Good

SUN

10

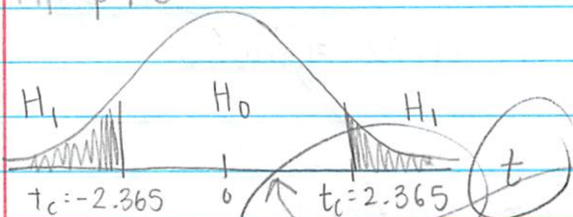
12.5 (2, 3, 6, 7)

2 a) $r \approx 0.18729$ conditions: a) set (x, y) is a random sample

$b \approx 0.17157$ b) normal distribution

b) $H_0: \rho = 0$ The average annual tuition and fees for 4 year public colleges and the avg annual full-time faculty salary have no correlation.

$H_1: \rho \neq 0$



$n = 9$ df = 7 $\alpha = 0.05$

$t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

$= \boxed{0.5045}$

we fail to rej~~ect H0~~ and reject H_1, $\alpha = 0.05$.

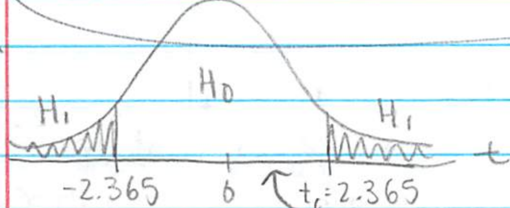
∃x insufficient statistical evidence to suggest that there is a correlation between the avg annual tuition and fees for 4 yr public colleges and the avg annual full-time faculty salary. In other words, there is no correlation.

P value ~~= 0.6294 > 0.05 (α)~~, weak evidence against the null

conditions c) $H_0: \beta = 0$ The population slope is zero.

a) random sample $H_1: \beta \neq 0$

b) normal distribution



~~n = 9~~ df = 7 $\alpha = 0.05$

$t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

$= \boxed{0.5045}$

we fail to reject ~~H0 and~~ reject H_0, $\alpha = 0.05$.

∃x insufficient statistical evidence to suggest that the slope isn't 0 between avg annual tuition/fees for 4yr public colleges and the avg annual full time faculty staff. In other words, slope is 0.

P value $= 0.6294 > 0.05 (\alpha)$, weak evidence against null

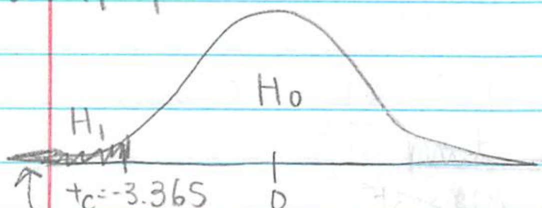
3a) $r = -0.97617$

$b = -0.05445$

conditions: random sample, nomal distribution

b) $H_0: \rho = 0$ There is no correlation between depth of dive (m) and optimal dive (hrs)

$H_1: \rho < 0$



$n = 7$ $df = 5$ $\alpha = 0.01$

$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \boxed{-10.06}$

we reject H_0 and fail to reject H_1, $\alpha = 0.01$

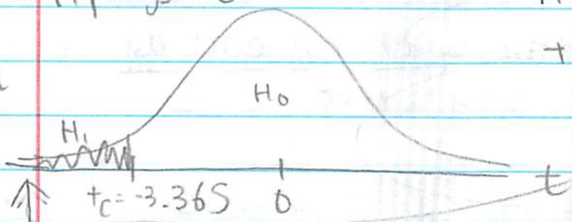
∃x sufficient statistical evidence to suggest that there is a negative correlation between depth of dive (m) and optimal time (hrs)

P value $= 8.31 \times 10^{-5} \approx 0 < 0.01 (\alpha)$, strong evidence against null

conditions: a) random sample b) normal distribution

c) $H_0: \beta = 0$ The pop slope is 0.

$H_1: \beta < 0$



$n = 7$ $df = 5$ $\alpha = 0.01$

$t = \boxed{-10.06}$

we reject H_0 and fail to reject H_1, $\alpha = 0.01$

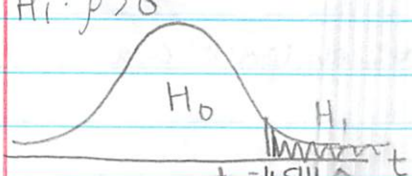
∃x sufficient statistical evidence to suggest that the slope is less than zero when considering the depth of a dive (m) and optimal time (hrs) for scuba divers.

6a) $r = 0.97694$ $b = 0.87937$

conditions: 1) set (x, y) is a random sample 2) normal distribution

b) $H_0: \rho = 0$ There is no correlation between list price for a random selection of cars of different models and options and the dealer invoice for the given vehicle.

$H_1: \rho > 0$



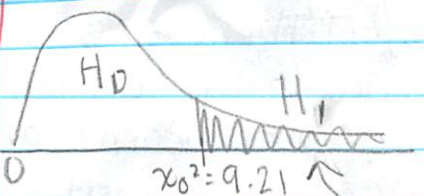
$n = 5$ $df = 3$ $\alpha = 0.01$

$t = \boxed{7.93}$

we reject H_0 and fail to reject H_1, $\alpha = 0.01$

∃x sufficient statistical evidence to suggest that there is a positive correlation between the list price for a random selection of cars of different models and the dealer invoice.

P value $= 0.0021 < \alpha (0.01)$, strong evidence against null



$df = E - 1 = 3 - 1 = 2$

$\alpha = 0.01$

$\chi_0^2 = 9.21$

$\chi^2 = \boxed{10.64}$

we reject H_0 and fail to reject H_1, $\alpha = 0.01$

∃X sufficient statistical evidence to suggest that the present buyers plan to keep their cars longer than buyers 10 yrs ago.

P value $= 0.0049 < 0.01(\alpha)$, strong evidence against null

7. a) $y = -3.69 + 0.9663x$

when $x = 9.5$, $y = 5.49$ thousand

b) $Se = 0.5368$

Interval:

$$E = t_c S_e \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_x}}$$

$$= 1.638(0.5368)\sqrt{1 + \frac{1}{5} + \frac{(9.5 - 8.72)^2}{3.028}}$$

$= 1.04$

$y_p - E < y < y_p + E$

$5.49 - 1.04 < y < 5.49 + 1.04$

$4.45\ y < 6.53$

If we repeated this sample 1000 times, we expect to capture the pop y values (per capita retail sales) @ 9,500 income 800 times. For this particular sample, we got an interval of $4.45 < y < 6.53$.

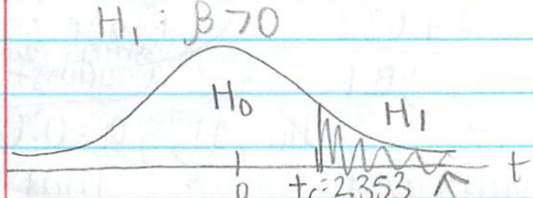
r =

c) $H_0: \rho = 0$ x and y have no correlation

b = 0.96631 d) $H_0: \beta = 0$ the pop slope is zero.

$H_1: \rho > 0$

$H_1: \beta > 0$



d.f. $= n - 2 = 5 - 2 = 3$

$t_0 = 2.353$

$t = \boxed{3.13}$

we reject Ho and fail to reject H_1, $\alpha = 0.05$
∃x sufficient statistical evidence to suggest that there is a corelation between per capita income and per capita retail sales and the slope is positive.
p value $= 0.02599 < 0.05(\alpha)$, strong evidence against null

8 a) $y = 0.633 + 0.694x$

$x = 5$, $y = 4.103$

b) $Se = 0.5506$

$$E = t_c Se \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{SSx}}$$

$$= 1.476(0.5506)\sqrt{1 + \frac{1}{7} + \frac{(5-3.47)^2}{12.614}}$$

$$= 0.9367$$

$$y_p - E < y < y_p + E$$

$$4.103 - 0.9367 < y < 4.103 + 0.9367$$

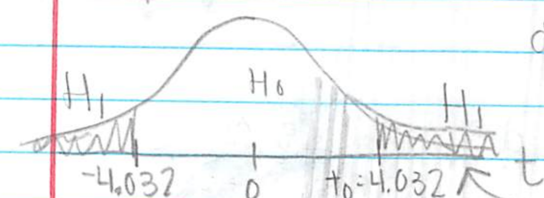
$$3.1663 < y < 5.0397$$

If we took this sample 1000 times, we expect to capture the pop y values of % change in consumer prices for the past yr in Australia, Austria, Canada, France, Italy, Spain & U.S. 800 times. For this particular sample, we got an interval of $3.1663 < y < 5.0397$

c) $r = 0.8947$

d) $b = 0.6943$

$H_0: \rho = 0$ x and y have no correlation
$H_0: \beta = 0$ The pop slope is zero.
$H_1: \rho \neq 0$
$H_1: \beta \neq 0$



d.f. $= n-2 = 7-2 = 5$

$t_0 = \pm 4.032$

$t = \boxed{4.48}$

France, Canada, Italy, U.S. have positive correlation and slope.
Pvalue $= 0.0065 < 0.01(\alpha)$
strong evidence against null

Ho, H1, $\alpha = 0.01$

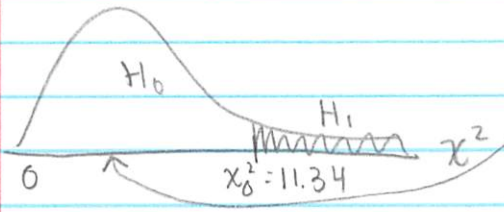
∃x sufficient statistical evidence to suggest that % chan in wages and % change in prices past year in Australia, Austria

SUN

12 RVW (1, 2, 4, 7, 8)

1. $H_0 : \chi^2 = 0$ The time to do a test and test results are independent.

$H_0 : \chi^2 > 0$



d.f. = (R−1)(C−1) = (1)(3) = 3

$\chi^2 = \boxed{3.92}$

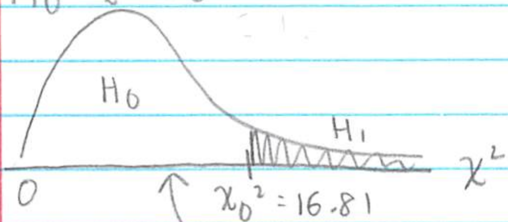
we fail to reject Ho and reject H_1.

∃x insufficient statistical evidence to suggest that the time to do a test and the test results are dependent. Thus, they are independent.

P value = 0.2697 > 0.01 (α), weak evidence against null

2. $H_0 : \chi^2 = 0$ Teacher ratings and student grades are independent.

$H_0 : \chi^2 > 0$



d.f. = (R−1)(C−1) = 2 x 3 = 6

$\chi_0^2 = 16.81$

$\chi^2 = \boxed{9.79}$

we fail to reject Ho and reject H_1, $\alpha = 0.01$

∃x insufficient statistical evidence to suggest that the teacher ratings and student grades are dependent. Thus, they are independent.

P value = 0.1337 > 0.01 (α), weak evidence against null

4.

Length (Time)	10 yrs Ago	O Now	E
> 5 yrs	20%	48	0.2 x 200 = 40
2 < yrs < 5	30%	75	0.3 x 200 = 60
< 2 yrs	50%	77	0.5 x 200 = 100

$H_0 : \chi^2 = 0$ Population of present car buyers fit distribution from 10 yrs ago.

$H_1 : \chi^2 > 0$

SUN

10-

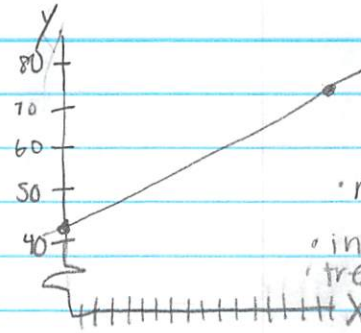
Holmes

36 a) $y = 43.66 + 1.6455x$

b) $x = 20$

$y = 76.57$ inches

x	y
18.08	73.42
0	43.66



Ana

- + association
- $r = 0.946$, high
- in/out
- trend: as length of stride increases, we see taller criminals
- $COD = r^2 = 0.8951$

90% of (v) in y (height) is explained by LSRL and (v) in x (stride length)

c) $E = t_c s_e \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{SS_x}}$

$= t_c s_e \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\Sigma x^2 + \frac{(\Sigma x)^2}{n}}}$

$= 2.228(2.81)\sqrt{1 + \frac{1}{12} + \frac{(20-18.08)^2}{4173 + \frac{(217)^2}{12}}}$

$= 6.52$

$y_p - E < y < y_p + E$

$76.57 - 6.52 < y < 76.57 + 6.52$

$70.01 < y < 83.13$

If we repeated this sample 1000 times, we expect to capture the pop height Y of a 20in stride person 950 times. For this particular sample, we got an interval of $70.05 < y < 83.09$ inches.

√5?

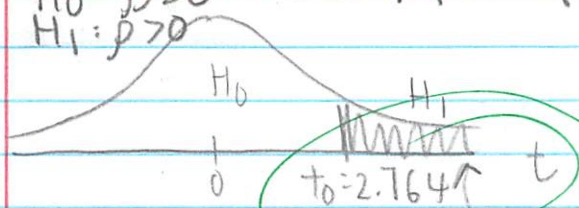
37 a) $r = 0.9461$

$H_0: \rho = 0$ x and y have no correlation.

$H_1: \rho > 0$

$H_0: \beta = 0$ the pop slope is 0.

$H_1: \beta > 0$



d.f. $= n-2 = 12-2 = 10$

$\alpha = 0.01$

$t_0 = 2.764$

$t = \boxed{9.2386}$

We reject H_0, fail to reject H_1, $\alpha = 0.01$ there is ∃X sufficient statistical evidence to suggest a correlation between length of stride and height and they have a positive slope.

p value $= 1.63 \times 10^{-6} \approx 0 < 0.01(\alpha)$, strong evidence against null

SUN

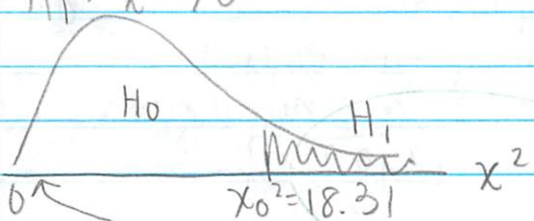
DUI

O	E
8	$0.037 \times 174 = 6.438$
35	$0.189 \times 174 = 32.886$
23	$0.129 \times 174 = 22.446$
19	$0.103 \times 174 = 17.922$
12	$0.085 \times 174 = 14.79$
14	$0.079 \times 174 = 13.746$
16	$0.08 \times 174 = 13.92$
13	$0.079 \times 174 = 13.746$
10	$0.068 \times 174 = 11.832$
9	$0.057 \times 174 = 9.918$
15	$0.094 \times 174 = 16.356$

a) $H_0: \chi^2 = 0$ The age distribution in Freemont county follows national distribution.

$H_1: \chi^2 > 0$

d)



$df = E - 1 = 11 - 1 = 10$

$\alpha = 0.05$

$\chi_0^2 = 18.31$

b) $\chi^2 = \boxed{1.9562}$

c) $E - 1 = 11 - 1 = 10$

e) We fail to reject H_0 and reject H_1, $\alpha = 0.05$

∃x insufficient statistical evidence to suggest the Freemont county age distribution doesn't follow national distribution of drunk driving arrests. Thus, it follows distribution.

P value $= 0.9967 \gg 0.05(\alpha)$, weak evidence against null

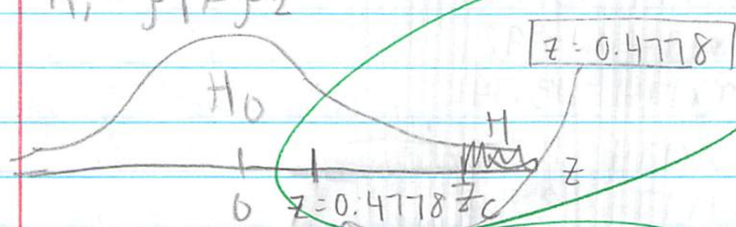
Ibuprofen

1. Given in ibuprofen group, $P = \frac{8}{672} = 0.0119$ that stomach is upset.

Given in placebo group, $P = \frac{6}{651} = 0.0092$ that stomach is upset.

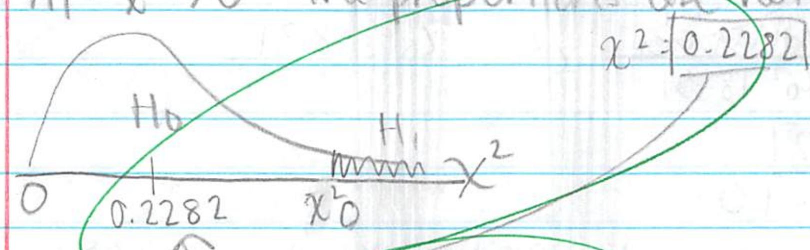
2. $H_0: \overset{\text{ibuprofen}}{p_1} = \overset{\text{placebo}}{p_2}$ there is no difference in proportions.

$H_1: p_1 \neq p_2$



P value = 0.6328 > 0.05, weak evidence against null
we fail to reject H_0 and reject H_1, $\alpha = 0.05$.
∃X insufficient statistical evidence to suggest proportion of stomach upsets is diff in ibuprofen and placebo group. Thus, they are the same

3. $H_0: \chi^2 = 0$ The proportions are equal.
$H_1: \chi^2 > 0$ the proportions are not equal



P value = 0.6328 > 0.05 (α), weak evidence against null
we fail to reject H_0, reject H_1, $\alpha = 0.05$
∃X insufficient statistical evidence to suggest proportions in ibuprofen and placebo groups are different. Thus, the proportions are the same.

4. The test conclusions agree. A test of 2 proportions can test if the proportions are not equal, less than, and more than each other more easily than χ^2 test.