STAT1302 – Statistical Analysis II
Assignment #2 – Solutions
Winter 2018

Q1. The management at New Century Bank claims that the mean waiting time for all customers at its branches is less than that at the Public Bank, which is its main competitor. A business consulting firm took a sample of 200 customers from the New Century Bank and found that they waited an average of 4.5 minutes before being served. Another sample of 300 customers taken from the Public Bank showed that these customers waited an average of 4.75 minutes before being served. Assume that the standard deviations for the two populations are 1.2 and 1.5 minutes, respectively.

a) Make a 97% confidence interval for the difference between the two population means.

- Let μ_1 and μ_2 be the population means waiting time for all customers at New Century Bank and at the Public Bank branches, respectively.
- From the given information, $n_1 = 200,\ \bar{x}_1 = 4.5,\ \sigma_1 = 1.2,\ n_2 = 300,\ \bar{x}_2 = 4.75,\ \sigma_2 = 1.5$
- The two samples are independent; The standard deviations $\sigma_1 = 1.2$ and $\sigma_2 = 1.5$ are known; Both sample sizes are large; so a 97% confidence interval for $\mu_1 - \mu_2$ is

$$(\bar{x}_1 - \bar{x}_2) \pm E, \text{ where } E = z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

- Point estimate of $\mu_1 - \mu_2$ is $\bar{x}_1 - \bar{x}_2 = 4.5 - 4.75 = -0.25$ min
- $1 - \alpha = 0.97 \rightarrow \alpha = 0.03 \rightarrow \alpha/2 = 0.015 \rightarrow z_{0.015} = 2.17$
- $E = z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 2.17 \times \sqrt{\frac{1.2^2}{200} + \frac{1.5^2}{300}} = 2.17 \times 0.12124 = 0.26$
- Therefore, a 97% confidence interval for $\mu_1 - \mu_2$ is $-0.25 \pm 0.26 \rightarrow (-0.51,\ 0.01)$
- We are 97% confident that the difference between the two population means waiting time for all customers at New Century Bank and the Public Bank branches is between -0.51 and 0.01 min.

b) Test at a 2.5% significance level whether the claim of the management of the New Century Bank is true. Use the critical value approach.

1) Identify the model parameters of interest. The parameters of interest are μ_1 and μ_2: μ_1 the population means waiting time for all customers at New Century Bank and at the Public Bank branches, respectively.

2) State the null and alternative hypotheses; H_0 and H_1. $\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 < 0 \end{cases}$

3) Select the distribution to use.
The two samples are independent; The standard deviations $\sigma_1 = 1.2$ and $\sigma_2 = 1.5$ are known; Both sample sizes are large; Therefore, we use the normal distribution to make the test.

4) Calculate the value of the test statistic and determine the rejection and non-rejection regions.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{-0.25 - 0}{0.12124} = -2.06$$

Left-tailed test: We reject H_0 if $z \leq -z_\alpha = -z_{0.025} = -1.96$

5) Make a decision. Under the critical value approach: $z = -2.06 < -1.96 \rightarrow$ Reject H_0

6) Conclusion: At a 2.5% level of significance, we have sufficient statistical evidence to conclude that the mean waiting time for all customers at New Century Bank is less than that at the Public, and the bank's claim is true.

c) Calculate the p-value for the test of part (b). Based on this p-value, would you reject the null hypothesis if $\alpha = 0.01$? What if $\alpha = 0.05$?

- Left-tailed test $\rightarrow P(Z < -2.06) = 0.0197$
- For $\alpha = 0.01$, do not reject H_0 since $0.0197 > 0.01$.
- For $\alpha = 0.05$, reject H_0 since $0.0197 < 0.05$.

Q2. A company sent seven of its employees to attend a course in building self-confidence. These employees were evaluated for their self-confidence before and after attending this course. The following table gives the scores (on a scale of 1 to 15, 1 being the lowest and 15 being the highest score) of these employees before and after they attended the course. Construct a 95% confidence interval for the difference between the mean of the scores of these employees before and the mean after they attended the course. State the underlying assumption(s).

Before	8	5	4	9	6	9	5
After	10	8	5	11	6	7	9
$\boldsymbol{d}$	-2	-3	-1	-2	0	2	-4

- $t_{0.025} = 2.447$ using $df = 7 - 1 = 6$
- $\bar{d} = -1.43, \quad \sum_{i=1}^{n} d_i^2 = 38, \quad s_d = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} d_i^2 - n\bar{d}^2\right)} = 1.9880$
- A 95% CI for μ_d: $\bar{d} \pm t_{0.025}\frac{s_d}{\sqrt{n}} = -1.43 \pm 1.84 \leftrightarrow (-3.27,\ 0.41)$
- We are 95% confident that the mean paired difference, the score of an employee before attending the course minus the score of the same employee after attending is between -3.27 and 0.41.

Q3. Quadro Corporation has two supermarket stores in a city. The company's quality control department wanted to check if the customers are equally satisfied with the service provided at these two stores. A sample of 380 customers selected from Supermarket I produced a mean satisfaction index of 7.6 (on a scale of 1 to 10, 1 being the lowest and 10 being the highest) with a variance of 0.5625. Another sample of 370 customers selected from Supermarket II produced a mean satisfaction index of 8.1 with a variance of 0.3481. Assume that the customer satisfaction index for each supermarket has unknown but same population standard deviation.

a) Construct a 98% confidence interval for the difference between the mean satisfaction indexes for all customers for the two supermarkets.

- Population 1: customers of Supermarket I; Population 2: customers of Supermarket II
- $n_1 = 380,\ \bar{x}_1 = 7.6, s_1^2 = 0.5625, n_2 = 370,\ \bar{x}_2 = 8.1,\ s_2^2 = 0.3481,\ \alpha = 0.02 \leftrightarrow \frac{\alpha}{2} = 0.01$
- The two samples are independent; The standard deviations of the two populations are unknown, but assumed to be equal; Both sample sizes are large; so, we use the t distribution.
- The 98% confidence interval for $\mu_1 - \mu_2$ is $(\bar{x}_1 - \bar{x}_2) \pm E$ where $E = t_{0.01}\ s_{\bar{x}_1 - \bar{x}_2}$
- $df = n_1 + n_2 - 2 = 380 + 370 - 2 = 748 \rightarrow t_{0.01} \approx z_{0.01} = 2.33$
- $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(380-1)\times 0.5625 + (370-1)\times 0.3481}{748}} = 0.6758 \rightarrow s_{\bar{x}_1 - \bar{x}_2} = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.0493$
- $E = t_{0.01}\ s_{\bar{x}_1 - \bar{x}_2} = 2.33 \times 0.0493 = 0.12$
- The 98% confidence interval for $\mu_1 - \mu_2$ is $(\bar{x}_1 - \bar{x}_2) \pm E \leftrightarrow -0.5 \pm 0.12 \leftrightarrow (-0.62,\ -0.38)$
- We are 98% confident that the difference in the mean satisfaction indexes for all customers for the two supermarkets lies between -0.62 and -0.38.

b) Test at a 1% significance level whether the mean satisfaction indexes for all customers for the two supermarkets are different. . Use either the p-value approach or the critical value approach.

1) Identify the model parameters of interest. μ_1 and μ_2: the means satisfaction index for Supermarket I and Supermarket II, respectively.
2) State the null and alternative hypotheses; H_0 and H_1. $\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{cases}$
3) Select the distribution to use. The two samples are independent; The standard deviations of the two populations are unknown, but assumed to be equal; Both sample sizes are large; so, we use the t distribution.

4) Calculate the value of the test statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{-0.5 - 0}{0.0493} = -10.13$$

4a) Calculate the p-value: Two-tailed test $\rightarrow$ p-value $= 2P(T < -10.13) < 0.002$ for $df = \infty$ $(df = 1148)$
4b) Determine the Rejection Regions: Two-tailed test: Reject H_0 if $t \geq t_{0.005} = 2.576$ or $t \leq -2.576$

5) Make a decision.
Under the p-value approach: p-value $< 0.002 < 0.01 = \alpha \rightarrow$ Reject H_0
Under the critical value approach: $t = -10.13 < 2.576 \rightarrow$ Reject H_0

6) Conclusion:
At a 1% level of significance, we have sufficient statistical evidence to conclude that the mean satisfaction index for Supermarket I is different from the mean for Supermarket II.

Q4. In a random sample of 800 men aged 25 to 35 years, 24% said they live with one or both parents. In another sample of 850 women of the same age group, 18% said that they live with one or both parents. Construct a 95% confidence interval for the difference between the proportions of all men and all women aged 25 to 35 years who live with one or both parents.

- Let p_1 and p_1 be the proportions of all men and all women aged 25 to 35 years who live with one or both parents, respectively.
- $\hat{p}_1 = 0.24$ and $\hat{p}_2 = 0.18$
- $n_1 = 800 \geq 30, n_2 = 850 \geq 30, n_1p_1 = 800 \times 0.24 = 192 > 5$, $n_1q_1 = 800 \times 0.76 = 608 > 5$, $n_2p_2 = 850 \times 0.18 = 153 > 5$ and $n_2q_2 = 850 \times 0.82 = 697 > 5 \rightarrow$ So, we use the normal distribution.
- The point estimate of $p_1 - p_2 = \hat{p}_1 - \hat{p}_2 = 0.24 - 0.18 = 0.06$
- Therefore, a 95% confidence interval $p_1 - p_2$ is $(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$
- $z_{\alpha/2} = z_{0.025} = 1.96$
- $s_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \sqrt{\frac{0.24 \times 0.76}{800} + \frac{0.18 \times 0.82}{850}} = 0.02$
- $E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 1.96 \times 0.02 = 0.04$
- The 95% confidence interval $p_1 - p_2$: $0.06 \pm 0.04 \leftrightarrow (0.02, 0.1)$
- We are 95% confident that the difference between the two population proportions is between 0.02 and 0.1.

Q5. The owner of a mosquito-infested fishing camp in Alaska wants to test the effectiveness of two rival brands of mosquito repellents, X and Y. During the first month of the season, eight people are chosen at random from those guests who agree to take part in the experiment. For each of these guests, Brand X is randomly applied to one arm and Brand Y is applied to the other arm. These guests fish for 4 hours, then the owner counts the number of bites on each arm. The table below shows the number of bites on the arm with Brand X and those on the arm with Brand Y for each guest.

a) Test at a 5% significance level whether the effectiveness of Brand X is different from the effectiveness of Brand Y. State the underlying assumption(s). Use the critical value approach.

Guest	A	B	C	D	E	F	G	H
Brand X	12	23	18	36	8	27	22	32
Brand Y	9	20	21	27	6	18	15	25
d	3	3	−3	9	2	9	7	7

- $\bar{d} = \frac{37}{8} = 4.63, \quad \sum_{i=1}^{n} d_i^2 = 291, \quad s_d = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} d_i^2 - n\bar{d}^2\right)} = 4.1382$

1) Identify the model parameters of interest.

Let μ_1 and μ_2 be the means number of bites on the arm with Brand X and Brand Y, respectively. Then $\mu_d = \mu_1 - \mu_2$ is the mean of the differences between the number of bites on the arm with Brand X and Brand Y.

2) State the null and alternative hypotheses; H_0 and H_1. $\begin{cases} H_0: \mu_d = 0 \\ H_1: \mu_d \neq 0 \end{cases}$

3) Select the distribution to use.

i. The standard deviation σ_d of the population of paired differences is unknown.

ii. The sample size is small (i.e., $n = 8 < 30$), and we need to assume that the population of paired differences is approximately normally distributed .

so, we use the t distribution to make the test about μ_d.

4) Calculate the value of the test statistic and determine the Rejection and Non-rejection Regions

The observed t is: $t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{4.63 - 0}{4.1383/\sqrt{8}} = 3.165$

Two-tailed test: We reject H_0 if $t \geq t_{\alpha/2} = t_{0.025} = 2.365$ or $t \leq -t_{\alpha/2} = -2.365$

5) Make a decision.

Under the critical value approach: $t = 3.165 > 2.365 \rightarrow$ Reject H_0

6) Conclusion:

At a 5% level of significance, we have sufficient statistical evidence to conclude that the mean number of bites on the arm with Brand X and the mean number of bites on the arm with Brand Y are different for all such guests, so the effectiveness of Brand X is different from the effectiveness of Brand Y.

b) What would your decision in part (a) be if the probability of making a type I error were zero? Explain.

Type I error $= \alpha = 0 \rightarrow$ There is no rejection region $\rightarrow$ We fail to reject for $\alpha = 0$, so we cannot conclude that the mean number of bites on the arm with Brand X and the mean number of bites on the arm with Brand Y are different for all such guests.

Q6. A mail-order company has two warehouses, one on the East Coast and the second on the West Coast. The company's policy is to mail all orders placed with it within 72 hours. The company's quality control department checks quite often whether or not this policy is maintained at the two warehouses. A recently taken sample of 300 orders placed with the warehouse on the East Coast showed that 279 of them were mailed within 72 hours. Another sample of 400 orders placed with the warehouse on the West Coast showed that 364 of them were mailed within 72 hours.

a) Using a 2.5% significance level, can you conclude that the proportion of all orders placed at the warehouse on the West Coast that are mailed within 72 hours is lower than the corresponding proportion for the warehouse on the East Coast? Use the critical value approach.

- Population 1: warehouse on the East Coast and Population 2: warehouse on the West Coast
- $\hat{p}_1 = \frac{279}{300} = 0.93$ and $\hat{p}_2 = \frac{364}{400} = 0.91$
- $\bar{p} = \frac{x_1+x_2}{n_1+n_2} = \frac{279+364}{300+400} = 0.919$ and $\bar{q} = 1 - \bar{p} = 1 - 0.919 = 0.081$
- $s_{\hat{p}_1-\hat{p}_2} = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{0.919 \times 0.081\left(\frac{1}{300}+\frac{1}{400}\right)} = 0.0208$

1) Identify the model parameters of interest. p_1 and p_2: the proportion of all orders placed at the warehouse on the East Coast and the West Coast that are mailed within 72 hours, respectively.

2) State the null and alternative hypotheses; H_0 and H_1. $\begin{cases} H_0: p_1 - p_2 = 0 \ (\text{or } p_1 = p_2) \\ H_1: p_1 - p_2 > 0 \ (\text{or } p_1 > p_2) \end{cases}$

3) Select the distribution to use.
$n_1 = 300 \geq 30, n_2 = 400 \geq 30, n_1p_1 = 300 \times 0.93 = 279 > 5$, $n_1q_1 = 300 \times 0.07 = 21 > 5, n_2p_2 = 400 \times 0.91 = 364 > 5$ and $n_2q_2 = 400 \times 0.09 = 36 > 5 \rightarrow$ So, we use the normal distribution.

4) Calculate the value of the test statistic and determine the rejection and non-rejection regions

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{(0.93 - 0.91) - 0}{0.0208} = 0.96$$

Right-tailed test: We reject H_0 if $z \geq 1.96$

5) Make a decision. Under the critical value approach: $0.96 < 1.96 \rightarrow$ Fail to reject H_0

6) Conclusion. At a 2.5% level of significance, we do not have sufficient statistical evidence to conclude that the proportion of all orders placed at the warehouse on the East Coast that are mailed within 72 hours is higher than the corresponding proportion for the warehouse on the West Coast (or we do not have sufficient statistical evidence to conclude that the proportion of all orders placed at the warehouse on the West Coast that are mailed within 72 hours is lower than the proportion for the warehouse on the East Coast).

b) What is the type I error in (a)? Explain. What is the probability of making such an error

- The probability of making such an error occurs when a true null hypothesis is rejected. In this question, A Type I error occurs when the proportions of all orders placed at the warehouse on the East Coast and on the West Coast that are mailed within 72 hours are the same, but we conclude that the proportion of all orders placed at the warehouse on the East Coast that are mailed within 72 hours is higher than the corresponding proportion for the warehouse on the West Coast.
- The probability of making a Type I error is $\alpha = 0.025$.

c) Find the p-value for the test in part (a). p-value $= P(Z > 0.96) = 0.1685$

Q7.

a) Determine the value of χ^2 for 23 degrees of freedom and an area of 0.990 in the left tail of the chi-square distribution curve. 41.638

b) Find the value of χ^2 for 4 degrees of freedom and 0.005 area in the right tail of the chi-square distribution curve. 14.860

c) Find the value of χ^2 for $df = 6$ and 0.05 area in the left tail of the chi-square distribution curve. 1.635

Q8. Chance Corporation produces beauty products. Two years ago the quality control department at the company conducted a survey of users of one of the company's products. The survey revealed that 53% of the users said the product was excellent, 31% said it was satisfactory, 7% said it was unsatisfactory, and 9% had no opinion. Assume that these percentages were true for the population of all users of this product at that time. After this survey was conducted, the company redesigned this product. A recent survey of 800 users of the redesigned product conducted by the quality control department at the company showed that 495 of the users think the product is excellent, 255 think it is satisfactory, 35 think it is unsatisfactory, and 15 have no opinion. Is the percentage distribution of the opinions of users of the redesigned product different from the percentage distribution of users of this product before it was redesigned? Use $\alpha = 0.025$.

1) Identify the parameters of interest: The proportions of users, p_1, p_2, p_3 and p_4, who think the product is excellent, satisfactory, unsatisfactory and who have no opinion, respectively.

2) State the null and alternative hypotheses; H_0 and H_1.

H_0: The percentage distribution of users' opinions is unchanged since the product was redesigned (or $p_1 = 0.53, p_2 = 0.31, p_3 = 0.07$ and $p_4 = 0.09$)

H_1: The percentage distribution of users' opinions has changed since the product was redesigned.

3) Select the distribution to use.

Since this is a multinomial experiment and $n = 800$ is large ($E_i > 5$), use the chi-square distribution.

4) Calculate the value of the test statistic and determine the rejection and non-rejection regions

Opinion	O	p	$E = np$	O^2	O^2/E
Excellent	495	0.53	424	245025	577.8892
Satisfactory	255	0.31	248	65025	262.2
Unsatisfactory	35	0.07	56	1225	21.875
No opinion	15	0.09	72	225	3.125
Total	$n = 800$	1	$n = 800$		865.087

$$\chi^2 = \sum \frac{O^2}{E} - n = 865.087 - 800 = 65.087$$

$k = 4$, $df = 4 - 1 = 3$ For $\alpha = 0.025$, the critical value of χ^2 is 9.348. We reject H_0 if $\chi^2 > 9.348$

5) Make a decision. $65.087 > 9.348 \rightarrow$ Reject H_0

6) Conclusion

At a 2.5% level of significance, we have sufficient statistical evidence to conclude that the percentage distribution of users' opinions has changed since the product was redesigned.

Q9. Four hundred people were selected from each of the four geographic regions (Midwest, Northeast, South, West) of the United States, and they were asked which form of camping they prefer. The choices were pop-up camper/trailer, family style (tenting with sanitary facilities), rustic (tenting, no sanitary facilities), or none. The results of the survey are shown in the following table. Based on the evidence from these samples, can you conclude that the distributions of favorite forms of camping are different for at least two of the regions? Use $\alpha = 0.01$.

1) Identify the characteristic of interest: The choices camper/trailer, family style, rustic, or none for four different geographic regions.
2) State the null and alternative hypotheses; H_0 and H_1.

$$\begin{cases} H_0: \text{The distributions of favorite forms of camping are the same for all four regions.} \\ H_1: \text{The distributions of favorite forms of camping are different for at least two regions.} \end{cases}$$

3) Select the distribution to use.
   Since this is a test of homogeneity and the sample size is large, $n = 1600$, use the chi-square distribution.
4) Calculate the value of the test statistic and determine the rejection and non-rejection regions

	Midwest	**Northeast**	**South**	**West**	**Total**
Camper/trailer	132 (131.25)	129 (131.25)	129 (131.25)	135 (131.25)	525
Family style	180 (167.25)	175 (167.25)	168 (167.25)	146 (167.25)	669
Rustic	46 (55.75)	50 (55.75)	59 (55.75)	68 (55.75)	223
None	42 (45.75)	46 (45.75)	44 (45.75)	51 (45.75)	183
Total	400	400	400	400	1600

$$\chi^2 = \sum \frac{O^2}{E} - n = \frac{132^2}{131.25} + \frac{129^2}{131.25} + \cdots + \frac{51^2}{45.75} - 1600 = 10.379$$

$df = (R-1)(C-1) = 3 \times 3 = 9$ For $\alpha = 0.01$, the critical value of χ^2 is 21.666. We reject H_0 if $\chi^2 > 21.666$

5) Make a decision. $10.739 < 21.666 \rightarrow$ Fail to reject H_0
6) Conclusion
   At a 1% level of significance, we do not have sufficient statistical evidence to conclude that the distributions of favorite forms of camping are different for at least two of the regions.

Q10. Many students graduate from college deeply in debt from student loans, credit card debts, and so on. A sociologist took a random sample of 401 single persons, classified them by gender, and asked, "Would you consider marrying someone who was \$25,000 or more in debt?" The results of this survey are shown in the following table. Test at a 1% significance level whether gender and response are related.

1) Identify the characteristics of interests: There are two characteristics of interest: gender and responses
2) State the null and alternative hypotheses; H_0 and H_1.

$$\begin{cases} H_0: \text{Gender and responses are independent.} \\ H_1: \text{Gender and responses are dependent.} \end{cases}$$

3) Select the distribution to use.
   Since this is a test of independence and $n = 401$ is large, use the chi-square distribution.

4) Calculate the value of the test statistic and determine the rejection and non-rejection regions

	Yes	No	Uncertain	Total
Women	125 (115.54)	59 (70.55)	21 (18.92)	205
Men	101 (110.46)	79 (67.45)	16 (18.08)	196
Total	226	138	37	401

$\chi^2 = \sum \frac{O^2}{E} - n = \frac{125^2}{115.54} + \frac{59^2}{70.55} + \cdots + \frac{16^2}{18.08} - 401 = 406.922 - 401 = 5.922$

$df = (R-1)(C-1) = 1 \times 2 = 2$ For $\alpha = 0.01$, we reject H_0 if $\chi^2 > 9.210$

5) Make a decision. $5.922 < 9.210 \rightarrow$ Fail to reject H_0

6) Conclusion. At a 1% level of significance, we do not have sufficient statistical evidence to conclude that gender and response are related.

Q11. In 2014, the variance of the ages of all workers at a large company that has more than 30,000 workers was 133. A recent random sample of 25 workers selected from this company showed that the variance of their ages is 112.

a) Using a 2.5% significance level, can you conclude that the current variance of the ages of workers at this company is lower than 133? Assume that the ages of all current workers at this company are (approximately) normally distributed.

1) Identify the model parameters of interest. σ^2: the variance of the ages of all workers at a large company.

2) State the null and alternative hypotheses; H_0 and H_1. $\begin{cases} H_0: \sigma^2 = 133 \\ H_1: \sigma^2 < 133 \end{cases}$

3) Select the distribution to use. Since this is a test about a population variance and the ages of all current workers at this company are (approximately) normally distributed, use the chi-square dist with $df = 24$.

4) Calculate the value of the test statistic and determine the rejection and non-rejection regions

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)112}{133} = 20.211$$

We reject H_0 if $0 < \chi^2 \leq 12.401$ $df = 25 - 1 = 24$

5) Make a decision. Since $20.211 > 12.401 \rightarrow$ Fail to reject H_0

6) Conclusion

At a 2.5% level of significance, we do not have sufficient statistical evidence to conclude that the current variance of the ages of workers at this company is lower than 133.

b) Construct a 98% confidence intervals for the variance of the ages of all current workers at this company.

χ^2 for 24 df and 0.01 area in the right tail $= 42.980$

χ^2 for 24 df and 0.99 area in the right tail $= 10.856$

The 98% confidence interval for σ^2 is $\frac{(n-1)s^2}{\chi^2_{\alpha/2}}$ to $\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} \rightarrow \frac{(25-1)112}{42.980}$ to $\frac{(25-1)112}{10.856} \rightarrow$ (62.5407,247.6050)