1. (13 marks = 2+3+2+1+2+1+1+1) From old records, it is believed that the average number of overtime hours for a company in Dammam is 20 hours per month. A random sample of 25 employees was selected and the number of overtime hours for each was registered. The sample gave an average of 24.75 and a standard deviation of 5.226. Do the data support the old records? Use the 2% significance level.

a. The test hypotheses are:

(1) $H_0 : \mu = 20$ vs $H_A : \mu \neq 20$ (1)

b. The assumptions are:

1. The population is normally distributed(1)

2. σ **is unknown**(1)

3. The sample size is small (1)

c. The critical value(s) is(are):

$$\pm t_{\frac{\alpha}{2},n-1} = \pm t_{\frac{0.02}{2},24} = \pm 2.492$$ (2)

d. The decision rule is:

Reject H_0 **if** $|T_{cal}| > t_{\frac{\alpha}{2},n-1} = 2.492$ **other wise don't reject** H_0 (1)

e. The test statistic is:

$$T_{cal} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(24.75 - 20)\sqrt{25}}{5.226} = 4.54458$$ (2)

f. The decision is:

Since $4.54458 > 2.492$**, reject the null hypothesis.** (1)

g. Your conclusion is:

There is not enough evidence to conclude that the average number of overtime hours for a company is 20 hours per month. (1)

h. Based on your decision, what type of error you might have committed?

Type I error (1)

2. (10 marks = 1+2+1+1+1+2+1+1) For the company in Dammam in the previous question (Question 1), test the hypothesis that the variance of the number of overtime hours for the company is less than 26.5 hours^2 (use 0.05 as significance level)

a. The estimate of the overall variance of the number of overtime hours for the company is:

$$S^2 = (5.226)^2 = 27.311076 \text{ (1)}$$

b. The test hypotheses are:

$$\text{(1) } H_0 : \sigma^2 \geq 26.5 \qquad vs \qquad H_A : \sigma^2 < 26.5 \text{ (1)}$$

c. The assumption is:

The population is normally distributed. (1)

d. The critical value(s) is (are):

$$\chi^2_{1-\alpha,n-1} = \chi^2_{1-0.05,25-1} = \chi^2_{0.95,24} = 13.84843 \text{ (1)}$$

e. The decision rule is:

Reject H_0 **if** $\chi^2_{cal} < \chi^2_{1-\alpha,n-1} = \chi^2_{0.95,24} = 13.84843$**, otherwise don't reject** H_0 (1)

f. The test statistic is:

$$\chi^2_{cal} = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(25-1)(5.226)^2}{26.5} = 24.7345594 \text{ (2)}$$

g. The decision is:

Since $24.7345594 > 13.84843$**, don't reject the null hypothesis.** (1)

h. Your conclusion is:

There is enough evidence to conclude that the variance of the number of overtime hours for a company is less than 26.5 hours^2. (1)

3. (14 marks =2+3+2+1+4+1+1) The Graystone Department Store study collected the following data on customer ages from random samples taken at two store locations:

Inner-City Store	**Suburban Store**
$n_1 = 19$	$n_2 = 11$
$\bar{x}_1 = 40$ years	$\bar{x}_2 = 35$ years
$s_1 = 9$ years	$s_2 = 10$ years

The information is used by the store to tailor product sales by customer ages at the two locations. At 5% significant level, test if there is no difference between the mean customer ages of Inner-City stores and the mean customer ages of Suburban stores.

a. The test hypotheses are:

(1) $H_0 : \mu_1 - \mu_2 = 0 \qquad vs \qquad H_A : \mu_1 - \mu_2 \neq 0$ (1)

b. The assumptions are:

1. **The populations are normally distributed**. (3)
2. σ_1 *and* σ_2 **are unknown.**
3. $\sigma_1 = \sigma_2$
4. **The samples are small.**
5. **The samples are independent.**

c. The critical value(s) is (are):

$$\pm t_{\frac{\alpha}{2}, n_1+n_2-2} = \pm t_{\frac{0.05}{2}, 19+11-2} = \pm t_{0.025,28} = \pm 2.048$$ (2)

d. The decision rule is:

Reject H_0 **if** $|T_{cal}| > t_{0.025,28} = 2.048$ **other wise don't reject** H_0 (1)

e. The test statistic is:

$$T_{cal} = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad where \quad s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(19-1)9^2 + (11-1)10^2}{19+11-2}$$ (2)

$$= \frac{40-35}{9.3694\sqrt{\frac{1}{19} + \frac{1}{11}}} \qquad = \frac{1458+1000}{28} = 87.7857 \Rightarrow s_p = 9.3694 = \mathbf{1.4085}$$ (2)

f. The decision is:

Since $1.4085 < 2.048$**, don't reject the null hypothesis.** (1)

g. Your conclusion is:

There is not enough evidence to conclude that there is a difference between the mean customer ages of Inner-City stores and the mean customer ages of Suburban stores. (1)

4. (9 marks = 2+1+1+1+2+1+1) In the previous question (question 3), test that there is no difference between the customers' age variance of Inner-City stores and of Suburban stores. Use 2% significance level.

a. The test hypotheses are:

(1) $H_0 : \sigma_1^2 = \sigma_2^2$ *vs* $H_A : \sigma_1^2 \neq \sigma_2^2$ (1)

b. The assumption is:

1. The populations are normally distributed (1)
2. The sample variances are independent

c. The critical value(s) is(are):

$$F_{\frac{\alpha}{2}, n_1 - 1, n_2 - 1} = F_{\frac{0.02}{2}, 11-1, 19-1} = F_{0.01,10,18} = 3.508 \text{ (1)}$$

d. The decision rule is:

Reject H_0 **if** $F_{cal} > F_{0.01,10,18} = 3.508$ **other wise don't reject** H_0 (1)

e. The test statistic is:

$$F_{cal} = \frac{S_1^2}{S_2^2} = \frac{10^2}{9^2} = \frac{100}{81} = 1.2345679 \text{ (2)}$$

f. The decision is:

Since $1.2346 < 3.508$**, don't reject the null hypothesis.** (1)

g. Your conclusion is:

There is not enough evidence to conclude that there is a difference between the customers' age variance of Inner-City stores and of Suburban stores. (1)

5. (11 marks = 1+2+2+2+2+1+1) Shell Oil office workers were asked which work schedule is most appealing: working five 8-hour days or four 10-hour days. A sample of 105 office workers showed that 67 **do not prefer** the four 10-hour day schedule. Do the given data suggest that Shell workers have no preference for one schedule over the other? Use the **p-value** approach with 1% significance level.

a. The sample proportion for the workers **preferring** four 10-hour days is:

$$\bar{p} = \frac{x}{n} = \frac{105-67}{105} = \frac{38}{105} = 0.3619 \text{ (1)}$$

b. The test hypotheses are:

$$\text{(1) } H_0 : p = 0.5 \quad vs \quad H_A : p \neq 0.5 \text{ (1)}$$

c. The assumptions are:

1. $n\bar{p} \geq 5 \Rightarrow 105(0.3619) = 38 > 5$ (1)

2. $n(1-\bar{p}) \geq 5 \Rightarrow 105(1-0.3619) = 67 > 5$ (1)

d. The test statistic is:

$$Z_{cal} = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.3619 - 0.5}{\sqrt{\frac{p0.5(1-0.5)}{105}}} = -2.8301 \text{ (2)}$$

e. The p-value is:

$$p - value = 2P(Z > |z_{cal}|) = 2P(Z > 2.83) = 2[0.5 - 0.4977] = 0.0046 \text{ (2)}$$

f. The decision is:

Reject H_0 **if** $p - value < \alpha = 0.01$ **other wise don't reject** H_0
Since $0.0046 < 0.01$**, reject the null hypothesis.** (1)

g. Your conclusion is:

There is enough evidence to conclude that Shell workers have a preference for one schedule over the other. (They don't like to work four 10-hour days). (1)

6. (13 marks=1+2+2+1+1+4+1+1) An accounting firm has been hired by a large computer company to determine whether the proportion of accounts receivables with errors in one division (Division 1) exceeds that of the second division (Division 2). The managers believe that such a difference may exist because of the loose standards employed by the first division. To conduct the test, the accounting firm has selected random samples of accounts from each division with the following results.

	Division 1	Division 2
Sample size	$n_1 = 100$	$n_2 = 100$
Number of errors found	$x_1 = 13$	$x_2 = 8$

Based on this information, use a significance level equal to 0.05 to help the accountants arrive at a decision. (Hint: you must answer each part below to arrive at the decision).

a. The pooled estimate for the overall proportion

$$\bar{p}_1 = \frac{x_1}{n_1} = \frac{13}{100} = 0.13, \quad \bar{p}_1 = \frac{x_1}{n_1} = \frac{8}{100} = 0.08 \text{ (2) to part f}$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{13+8}{100+100} = \frac{21}{200} = 0.105 \text{ (1)}$$

b. The test hypotheses are:

(1) $H_0 : p_1 - p_2 \le 0 \quad vs \quad H_A : p_1 - p_2 > 0$ (1)

c. The assumption(s) is(are):

1. $n_1\bar{p}_1 \ge 5 \Rightarrow 100(0.13) = 13 > 5$ **and** $n_1(1-\bar{p}_1) \ge 5 \Rightarrow 100(1-0.13) = 87 > 5$ (1)
2. $n_2\bar{p}_2 \ge 5 \Rightarrow 100(0.08) = 8 > 5$ **and** $n_2(1-\bar{p}_2) \ge 5 \Rightarrow 100(1-0.08) = 92 > 5$ (1)

d. The critical value(s) is(are): $Z_\alpha = Z_{0.05} = 1.645$ (1)

e. The decision rule is:

Reject H_0 **if** $Z_{cal} > Z_{0.05} = 1.645$ **other wise don't reject** H_0 (1)

f. The test statistic is:

$$Z_{cal} = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.13 - 0.08}{\sqrt{0.105(1-0.105)\left(\frac{1}{100} + \frac{1}{100}\right)}} = \frac{0.05}{2.16766} = 0.02306 \text{ (2)}$$

g. The decision is:

Since $0.02306 < 1.645$, **don't reject the null hypothesis.** (1)

h. Your conclusion is:

There is no enough evidence to determine whether the proportion of accounts receivables with errors in Division 1 exceeds that of the second division (Division 2). (1)