THE UNIVERSITY OF SYDNEY

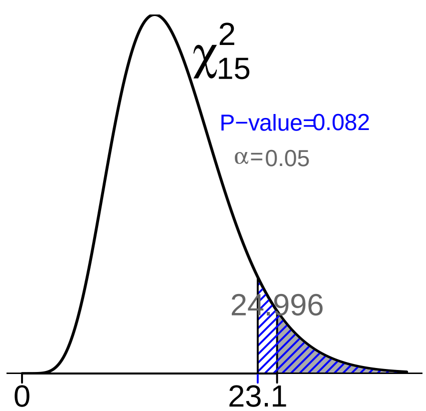
STAT2012 STATISTICAL TESTS

Semester 2	**Solution to Tutorial Week 6**	2015

Tutorial questions

1. (a) The Chi-square test for the variance of the closing prices of the first stock is

   1. **Hypothesis:** $H_0: \ \sigma^2 = 1 \quad \text{vs} \quad H_1 : \sigma^2 > 1.$
   2. **Test statistic:** $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2} = \dfrac{15 \times 1.54}{1} = 23.1$
   3. **Assumption:** $X_i \sim \mathcal{N}(\mu, 1)$ under H_0. Then $\chi_0^2 \sim \chi_{n-1}^2$.
   4. ***P*-value:** $\Pr(\chi_{15}^2 \geq 23.1) \in (0.05, 0.1)$ ($\chi_{15,0.9}^2 = 22.3,\ \chi_{15,0.95}^2 = 25.0;\ 0.082$ from R)
   5. **Decision:** Since the p-value > 0.05, the data is consistent with H_0 that the variance of the closing prices of the first stock is 1.



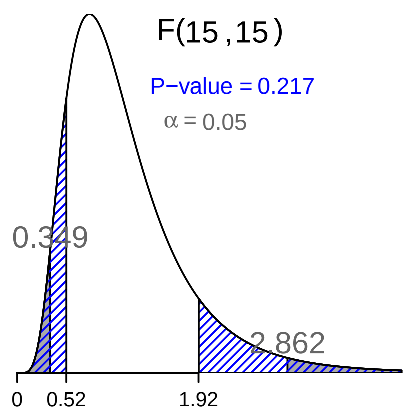
   (b) The F test for the difference in variances of the closing prices of the two stocks is

   1. **Hypotheses:** $H_0 : \ \sigma_x^2 = \sigma_y^2 \quad \text{vs} \ H_1 : \ \sigma_x^2 \neq \sigma_y^2$
   2. **Test statistic:** $f_0 = s_y^2/s_x^2 = \dfrac{2.96}{1.54} = 1.922.$
   3. **Assumption:** $X_i \sim \mathcal{N}(\mu_x, \sigma_x^2)$, $Y_i \sim \mathcal{N}(\mu_y, \sigma_y^2)$ and X_i & Y_i are independent.
   4. ***P*-value:** p-value $= 2\Pr(F_{15,15} \geq 1.922) = 2(0.1086) = 0.2173$ (from R),
      or p-value $= 2\Pr(F_{15,15} \geq 1.922) > 2(0.1) > 0.2$ since $F_{15,15,0.9} = 1.97$.
   5. **Decision:** The data is consistent with H_0. There is no evidence to indicate a difference in the variability of closing prices of two stocks.

(c) The 95% confidence interval for the ratio of variances $\frac{\sigma_y^2}{\sigma_x^2}$:

$$\begin{aligned}
& \left(\frac{s_y^2}{s_x^2}/F_{n_y-1,n_x-1,1-\alpha/2},\ \frac{s_y^2}{s_x^2}/F_{n_y-1,n_x-1,\alpha/2}\right) \\
= & \left(\frac{2.96}{1.54}/F_{15,15,0.975},\ \frac{2.96}{1.54}/F_{15,15,0.025}\right) \\
= & \left(\frac{2.96}{1.54}/F_{15,15,0.975},\ \frac{2.96}{1.54}/\frac{1}{F_{15,15,0.975}}\right) \text{ (lower to upper area in table; swap df \& take reciprocal)} \\
= & \left(\frac{2.96}{1.54}/2.86,\ \frac{2.96}{1.54}/\frac{1}{2.86}\right) = \left(\frac{2.96}{1.54}/2.86,\ \frac{2.96}{1.54}\times 2.86\right) \\
= & (0.672,\ 5.501)
\end{aligned}$$

Since the hypothesized ratio of $\frac{\sigma_y^2}{\sigma_x^2} = 1$ lies in the 95% CI, the data are consistent with H_0. The variances of the closing prices of the two stocks are not significantly different.



2. Let $\mu_i, i = 1, 2, 3, 4$, denote the mean increase in heartbeats per minute for the groups 10-19, 20-39, 40-59 and 60-69 respectively. The groups information is

n_i	9	10	10	10	$N = 39$	$g = 4$
$\bar{y}_{i.}$	30	27.5	29.5	28.2		

$$CM = N\bar{y}^2 = 32279.08, \qquad \sum\sum y_{ij}^2 = 33180.$$

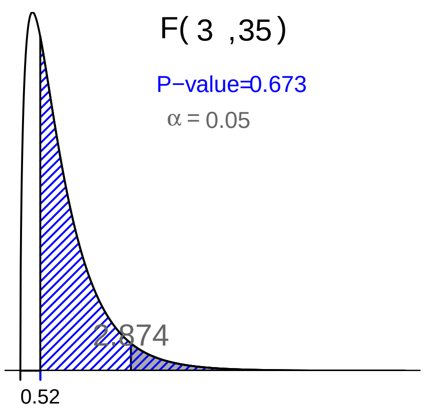
The one-way ANOVA test for the difference in mean increase in heart beats across the four age groups is

1. **Hypotheses:** $H_0: \mu_1 = ... = \mu_4$ vs
   $H_1:$ at least one of the equalities does not hold.
2. **Test statistic:**

$$f_0 = \frac{MST}{MSR} = \frac{\left[(\sum_{i=1}^{g} n_i\bar{y}_i^2 - N\bar{y}^2\right]/(g-1)}{\sum_{i=1}^{g}(n_i-1)s_i^2/(N-g)} = \frac{38.32/3}{862.6/35} = \frac{12.77}{24.645} = 0.518.$$

$$
\begin{aligned}
CM &= N\bar{y}^2 = 32279.08 \\
SST_0 &= \sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}^2 - N\bar{y}^2 = 33180 - 32279.08 = 900.92 \\
SST &= \sum_{i=1}^{g} n_i(\bar{y}_{i.})^2 - N\bar{y}^2 \\
&= [9(30^2) + 10(27.5^2) + 10(29.5^2) + 10(28.2^2)] - 32279.08 \\
&= 38.32 \\
SSR &= 900.92 - 38.32 = 862.6
\end{aligned}
$$

3. **Assumption:** $Y_{ij} \sim \mathcal{N}(\mu_i, \sigma^2)$ and Y_{ij} & $Y_{i'j'}$ are independent. Then $f_0 \sim F_{g-1,N-g}$.
4. P**-value:** $\Pr(F_{3,35} \geq 0.518) = 0.6725 > 0.10$ ($F_{3,30,0.900} = 2.28$; 0.6726 from R).
5. **Decision:** The data is consistent with the null hypothesis, i.e., there is no evidence to indicate a difference in the mean increase in heart beats across the four age groups.



ANOVA table

Source	df	SS	MS	F
Groups	3	38.32	12.77	0.518
Residuals	35	862.6	24.645	
Total	38	900.92		

Since the test for the equality of means is not significant, the means of the increase in heartbeats per minute for the age groups 10-19 and 20-39 should be the same.

Extra problems

1. The 95% upper sided confidence interval for the true of variance in question 1(a) is

$$\left(s^2\frac{n-1}{\chi^2_{n-1,1-\alpha}},\ \infty\right) = \left(1.54\,\frac{16-1}{\chi^2_{15,0.95}},\ \infty\right) = \left(1.54\,\frac{16-1}{24.996},\ \infty\right)$$
$$= (0.9241479,\ \infty)$$

Since the hypothesized variance of $\sigma^2 = 1$ is included in the 95% CI, the data are consistent with H_0. The variance of the closing prices of the first stock is not significantly greater than 1.

2. (a) The F test for the difference in variances of the bacteria counts to two methods of irradiation is

   1. **Hypotheses:** $H_0:\ \sigma_x^2 = \sigma_y^2$ vs $H_1:\ \sigma_x^2 \neq \sigma_y^2$
   2. **Test statistic:** $f_0 = s_y^2/s_x^2 = \frac{841}{324} = 2.5877$.
   3. **Assumption:** $X_i \sim \mathcal{N}(\mu_x, \sigma_x^2)$, $Y_i \sim \mathcal{N}(\mu_y, \sigma_y^2)$ and X_i & Y_i are independent.
   4. **P-value:** p-value $= 2\Pr(F_{20,30} \geq 2.5877) \in (2(0.005), 2(0.01)) = (0.01, 0.02)$ or $2(0.0091) = 0.0181$ (from R), since $F_{20,30,0.99} = 2.55$ and $F_{20,30,0.995} = 2.82$.
   5. **Decision:** Since p-value < 0.05. There is strong evidence in the data that the variability of the bacteria counts for the two methods of irradiation differ.

   (b) The 95% confidence interval for the ratio of the variances of the number of re-recordings of the two types of tape is

$$\left(\frac{s_y^2}{s_x^2}/F_{n_y-1,n_x-1,1-\alpha/2},\ \frac{s_y^2}{s_x^2}/F_{n_y-1,n_x-1,\alpha/2}\right)$$
$$= \left(\frac{841}{324}/F_{20,30,0.975},\ \frac{841}{324}/F_{20,30,0.025}\right)$$
$$= \left(\frac{841}{324}/F_{20,30,0.975},\ \frac{841}{324}/\frac{1}{F_{30,20,0.975}}\right) \text{ (change lower area to upper area in table)}$$
$$= \left(\frac{841}{324}/2.20,\ \frac{841}{324}/\frac{1}{2.35}\right) = \left(\frac{841}{324}/2.20,\ \frac{841}{324}\times 2.35\right)$$
$$= (1.1654,\ 6.099846)$$

   We estimate σ_y^2/σ_x^2 to be between 1.1654 and 6.0998. The F values are obtained from F table.

3. Let $\mu_i, i = 1, 2, 3$ denote the mean strength for the three alloys respectively. The groups information is

n_i	4	6	5	$N = 15$ $g = 3$
$\bar{y}_{i.}$	20.75	21.5	26	$\bar{y} = 22.8$

The one-way ANOVA test for the difference in mean increase in heart beats across the four age groups is

1. **Hypotheses:** $H_0:\ \mu_1 = \mu_2 = \mu_3$ vs
   $H_1:$ at least one of the equalities does not hold.

2. **Test statistic:** $f_0 = \dfrac{MST}{MSR} = \dfrac{\left(\sum_{i=1}^{g} n_i \bar{y}_i^2 - N\bar{y}^2\right)/(g-1)}{\sum_{i=1}^{g}(n_i-1)s_i^2/(N-g)} = \dfrac{39.017}{21.354} = 1.8299.$

$$\begin{aligned}
CM &= n\bar{y}^2 = 15(22.8^2) = 7797.6 \\
SST_0 &= \sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}^2 - n\bar{y}^2 = 8132 - 7797.6 = 334.4 \\
SST &= \sum_{i=1}^{g} n_i(\bar{y}_{i.})^2 - N\bar{y}^2 \\
&= [4(20.75^2) + 6(27.5^2) + 5(29.5^2)] - 7797.6 \\
&= 78.15 \\
SSR &= 334.4 - 78.15 = 256.25
\end{aligned}$$

3. **Assumption:** $Y_{ij} \sim \mathcal{N}(\mu_i, \sigma^2)$ and Y_{ij} & $Y_{i'j'}$ are independent. Then $f_0 \sim F_{g-1,N-g}$.
4. ***P*-value:** p-value $= \Pr(F_{2,12} \geq 1.8299) = 0.2025$ (from R),
   or p-value > 0.1 (since $F_{2,12,0.90} = 2.81$).
5. **Decision:** The data is consistent with the null hypothesis, i.e., there is no evidence to indicate a difference in the mean strengths of the welds with different alloys.

ANOVA table

Source	df	SS	MS	F	p
Groups	2	78.15	39.017	1.8299	0.2025
Residuals	12	256.25	21.354		
Total	14	334.4			