Assignment Four

① (a) $Y \sim N(a\mu + b, a^2\sigma^2)$

(b) $Z \sim N(0,1)$

(c) $Y \sim N(n\mu, n\sigma^2)$

(d) $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

(e) $Z \sim N(0,1)$

(f) $Y \sim N(\sum_{i=1}^{n} a_i\mu_i , \sum_{i=1}^{n} a_i^2\sigma_i^2)$

② (a) $M_Y(t) = M_{\sum_{i=1}^{n} X_i}(t) \overset{\text{ind}}{=} \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} (1-2t)^{-\nu_i/2}$

$= (1-2t)^{-(\sum_{i=1}^{n}\nu_i)/2}$ MGF of $\chi^2(\sum_{i=1}^{n}\nu_i)$

(b) $Y = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$, because

- $Z_i = \frac{X_i - \mu}{\sigma} \sim N(0,1)$
- $Z_i^2 \sim \chi^2(1)$
- Z_i are independent because functions of independent random variables are independent.
- Sum of independent chi-squares is chi-squared, with df = sum of df. This is Q 2a.

(c) By independence, $M_Y(t) = M_{Y_1}(t) M_{Y_2}(t)$

Because Y & Y_1 are chi-squared,

$(1-2t)^{-(\nu_1+\nu_2)/2} = (1-2t)^{-\nu_1/2} M_{Y_2}(t)$

$\Rightarrow \frac{\cancel{(1-2t)^{-\nu_1/2}}(1-2t)^{-\nu_2/2}}{\cancel{(1-2t)^{-\nu_1/2}}} = M_{Y_2}(t)$ MGF of $\chi^2(\nu_2)$

3 (a) $a_j = 1 - \frac{1}{n}$, $a_i = -\frac{1}{n}$ for $i \neq j$

(b) $\text{Cov}(\bar{X}, X_j - \bar{X}) = \text{Cov}(\bar{X}, X_j) - \text{Cov}(\bar{X}, \bar{X})$

$$= \text{Cov}\left(X_j, \frac{1}{n}\sum_{i=1}^{n} X_i\right) - \text{Var}(\bar{X})$$

$$= \frac{1}{n}\sum_{i=1}^{n} \text{Cov}(X_i, X_j) - \frac{\sigma^2}{n}$$

$$= \frac{1}{n}\left[\text{Cov}(X_j, X_j) + \sum_{i \neq j} \text{Cov}(X_i, X_j)\right] - \frac{\sigma^2}{n}$$

$$\overset{\text{ind}}{=} \frac{1}{n}\left[\text{Cov}(X_j, X_j) + 0\right] - \frac{\sigma^2}{n}$$

$$= \frac{1}{n}\text{Var}(X_j) - \frac{\sigma^2}{n} = \frac{1}{n}\sigma^2 - \frac{\sigma^2}{n} = 0$$

(Does not depend on normality)

(c) $\bar{X}$ and $X_j - \bar{X}$ are both linear combinations of $X_1 \ldots X_n$, as in Q3a above. So as in Q1f, they are normally distributed. Since their covariance is zero by Q3b and zero covariance implies independence for the normal, $\bar{X}$ is independent of $X_j - \bar{X}$ for $j = 1, \ldots, n$. Finally S^2 is a function of $(X_1 - \bar{X}), \ldots, (X_n - \bar{X})$, so because functions of independent random variables are independent, S^2 is independent of $\bar{X}$.

This answer is more detailed than you need. Could just say: $\text{Cov}(\bar{X}, X_j - \bar{X}) = 0$, and they are normal so they are independent. Functions of independent and S^2 is a function of $(X_1 - \bar{X}), \ldots, (X_n - \bar{X})$, so it is independent of $\bar{X}$.

(3d) $\sum_{i=1}^{n}(X_i-\mu)^2 = \sum_{i=1}^{n}(X_i-\bar{X}+\bar{X}-\mu)^2$

$= \sum_{i=1}^{n}\left[(X_i-\bar{X})^2 + 2(X_i-\bar{X})(\bar{X}-\mu) + (\bar{X}-\mu)^2\right]$

$= \sum_{i=1}^{n}(X_i-\bar{X})^2 + 2(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\bar{X}) + \sum_{i=1}^{n}(\bar{X}-\mu)^2$

$= \sum_{i=1}^{n}(X_i-\bar{X})^2 + 2(\bar{X}-\mu)\left(\sum_{i=1}^{n}X_i - \sum_{i=1}^{n}\bar{X}\right) + n(\bar{X}-\mu)^2$

$= \sum_{i=1}^{n}(X_i-\bar{X})^2 + 2(\bar{X}-\mu)(\Sigma X_i - n\bar{X}) + n(\bar{X}-\mu)^2$

$= \sum_{i=1}^{n}(X_i-\bar{X})^2 + 2(\bar{X}-\mu)\underbrace{\left(\Sigma X_i - n\frac{\Sigma X_i}{n}\right)}_{=0} + n(\bar{X}-\mu)^2$

$= \sum_{i=1}^{n}(X_i-\bar{X})^2 + n(\bar{X}-\mu)^2$

(e) From (3d), $\sum_{i=1}^{n}(X_i-\mu)^2 = \sum_{i=1}^{n}(X_i-\bar{X})^2 + n(\bar{X}-\mu)^2$

$\Longrightarrow \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} + \frac{n(\bar{X}-\mu)^2}{\sigma^2}$

$\Longrightarrow \underbrace{\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2}_{= Y} = \underbrace{\frac{(n-1)S^2}{\sigma^2}}_{= Y_1} + \underbrace{\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2}_{= Y_2}$

$Y \sim \chi^2(n)$ by (2b), $Y_2 \sim N(0,1)$ by (1e)

Y_1 is a function of S^2, Y_2 is a function of $\bar{X}$ & S^2 and $\bar{X}$ are independent, so Y_1 & Y_2 are independent.

Then $Y_1 \sim \chi^2(n-1)$ by (Q2c).

(4) (a) $Z = \dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$, $Y = \dfrac{(n-1)S^2}{\sigma^2}$

Independent because Z is a function of $\bar{X}$ & Y is a function of S^2, and $\bar{X}$ & S^2 are independent

$$T = \frac{\dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/(n-1)}} = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t(n-1).$$

(b) $1-\alpha = P(-t_{1-\alpha/2} < T < t_{1-\alpha/2})$

$= P\left(-t_{1-\frac{\alpha}{2}} < \dfrac{\sqrt{n}(\bar{X}-\mu)}{S} < t_{1-\alpha/2}\right)$

$= P\left(-t_{1-\alpha/2}\dfrac{S}{\sqrt{n}} < \bar{X}-\mu < t_{1-\alpha/2}\dfrac{S}{\sqrt{n}}\right)$

$= P\left(-\bar{X} - t_{1-\alpha/2}\dfrac{S}{\sqrt{n}} < -\mu < -\bar{X} + t_{1-\alpha/2}\dfrac{S}{\sqrt{n}}\right)$

$= P\left(\bar{X} + t_{1-\alpha/2}\dfrac{S}{\sqrt{n}} > \mu > \bar{X} - t_{1-\alpha/2}\dfrac{S}{\sqrt{n}}\right)$

CI of $\left(\bar{X} - t_{1-\alpha/2}\dfrac{S}{\sqrt{n}}, \bar{X} + t_{1-\alpha/2}\dfrac{S}{\sqrt{n}}\right)$

or $\bar{X} \pm t_{1-\alpha/2}\dfrac{S}{\sqrt{n}}$

(4c) $\bar{x} \pm t_{1-\alpha/2} \frac{s}{\sqrt{n}} = -1.58 \pm 2.262 \frac{\sqrt{1.513}}{\sqrt{10}} = -1.58 \pm 0.88$

$= (-2.46, -0.70)$

(d) Yes. The confidence interval is all negative numbers, suggesting that Drug 2 worked better.

(e) $1-\alpha = P\left(\chi^2_{\alpha/2} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{1-\alpha/2}\right)$

$$= P\left(\frac{1}{\chi^2_{\alpha/2}} > \frac{\sigma^2}{(n-1)s^2} > \frac{1}{\chi^2_{1-\alpha/2}}\right)$$

$$= P\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}} > \sigma^2 > \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}\right)$$

$$= P\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\alpha/2}}\right)$$

(f) With $r=9$ df, $\chi^2_{0.025} = 2.70$ & $\chi^2_{0.975} = 19.02$

So the CI is

$$\left(\frac{(10-1)\,1.513}{19.02}, \frac{(10-1)\,1.513}{2.7}\right) = (0.72, 5.04)$$

5 (a) $\bar{X} - \bar{Y}$

(b) $\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\right)$

(c)

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

(d)

$$Y = \frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2)$$

They are independent because S_1^2 is a function of the X_i, S_2^2 is a function of the Y_i, & the X_i and Y_i are independent.

(e)

$$T = \frac{\dfrac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\cancel{\sigma}\sqrt{1/n_1 + 1/n_2}}}{\sqrt{\dfrac{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\cancel{\sigma}^2}}{n_1 + n_2 - 2}}} = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}\sqrt{1/n_1 + 1/n_2}}$$

(f) $$= \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}}$$

(g) $df = n_1 + n_2 - 2$

(5h) $1-\alpha = P\left(-t_{1-\frac{\alpha}{2}} < \dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_p\sqrt{1/n_1+1/n_2}} < t_{1-\frac{\alpha}{2}}\right)$

$$\overset{\text{2 steps}}{=} P\left(-t_{1-\alpha/2} < \frac{\mu_1-\mu_2-(\bar{X}-\bar{Y})}{S_p\sqrt{1/n_1+1/n_2}} < t_{1-\alpha/2}\right)$$

$$= P\left(-t_{1-\alpha/2} S_p\sqrt{\tfrac{1}{n_1}+\tfrac{1}{n_2}} < \mu_1-\mu_2-(\bar{X}-\bar{Y}) < t_{1-\alpha/2} S_p\sqrt{\tfrac{1}{n_1}+\tfrac{1}{n_2}}\right)$$

$$= P\left(\bar{X}-\bar{Y}-t_{1-\frac{\alpha}{2}} S_p\sqrt{\tfrac{1}{n_1}+\tfrac{1}{n_2}} < \mu_1-\mu_2 < \bar{X}-\bar{Y}+t_{1-\alpha/2} S_p\sqrt{\tfrac{1}{n_1}+\tfrac{1}{n_2}}\right)$$

(i) $s_p = \sqrt{\dfrac{(9-1)48.2+(7-1)32.7}{9+7-2}} = 6.45,$

$$\bar{x}-\bar{y} \pm t_{1-\frac{\alpha}{2}}\, s_p\sqrt{\tfrac{1}{n_1}+\tfrac{1}{n_2}}$$

$$= 14.1-13.3 \pm 2.145 * 6.45\sqrt{\tfrac{1}{9}+\tfrac{1}{7}}$$

$$= 0.8 \pm 6.97 \quad = (-6.17,\ 7.77)$$

(j) Nobody. Zero difference is within the confidence interval.

6 (a) $[0,\infty)$ Open or closed does not matter

(b) i) $F_1 = \dfrac{Y_1/\nu_1}{Y_2/\nu_2}$, $F_2 = \dfrac{1}{F_1} = \dfrac{Y_2/\nu_2}{Y_1/\nu_1} \sim F(\nu_2, \nu_1)$

ii) $P(F_1 \le x) = P\left(\dfrac{1}{F_1} \ge \dfrac{1}{x}\right) = P\left(F_2 \ge \dfrac{1}{x}\right)$

(c)

$$F = \frac{\dfrac{(n_1-1)S_1^2}{\sigma_1^2}\Big/(n_1-1)}{\dfrac{(n_2-1)S_2^2}{\sigma_2^2}\Big/(n_2-1)} = \frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}$$

$$\sim F(n_1-1, n_2-1)$$

Since we want a CI for $\dfrac{\sigma_1^2}{\sigma_2^2}$, a better pivotal quantity is $\dfrac{S_2^2}{S_1^2}\,\dfrac{\sigma_1^2}{\sigma_2^2} \sim F(n_2-1, n_1-1)$

by (b) part i)

(d) $1-\alpha = P\left(f_{\alpha/2} < \dfrac{S_2^2}{S_1^2}\,\dfrac{\sigma_1^2}{\sigma_2^2} < f_{1-\alpha/2}\right)$

$= P\left(f_{\alpha/2}\,\dfrac{S_1^2}{S_2^2} < \dfrac{\sigma_1^2}{\sigma_2^2} < f_{1-\frac{\alpha}{2}}\,\dfrac{S_1^2}{S_2^2}\right)$ with df $= 6, 8$

(e) $\left(0.1786\left(\dfrac{48.2}{32.7}\right),\ 4.65\left(\dfrac{48.2}{32.7}\right)\right) = (0.26,\ 6.85)$

(f) No. $\dfrac{\sigma_1^2}{\sigma_2^2} = 1$ is in the confidence interval.