Sec. 7.2 Ex. 3

We have two independent random samples $\begin{cases} X_1, \ldots, X_{m_x} \sim N(\mu_x, \sigma_x^2) \text{ with } m_x = 12 \\ Y_1, \ldots, Y_{m_y} \sim N(\mu_y, \sigma_y^2) \text{ with } m_y = 15 \end{cases}$

with unknown but equal variances $\sigma_x^2 = \sigma_y^2 = \sigma^2$

Recall that:

$$\bar{X} \sim N\left(\mu_x, \frac{\sigma^2}{m_x}\right) \text{ and } \bar{Y} \sim N\left(\mu_y, \frac{\sigma^2}{m_y}\right) \text{ and so } \bar{X} - \bar{Y} \sim N\left(\mu_x - \mu_y, \frac{\sigma^2}{m_x} + \frac{\sigma^2}{m_y}\right)$$

Thus, we have that $Z = \dfrac{\bar{X} - \bar{Y} - (\mu_x - \mu_y)}{\sqrt{\dfrac{\sigma^2}{m_x} + \dfrac{\sigma^2}{m_y}}} \sim N(0,1)$ and so, from

Recall also that $\dfrac{(m_x-1)S_x^2}{\sigma^2} \sim \chi^2(m_x-1)$ and $\dfrac{(m_y-1)S_y^2}{\sigma^2} \sim \chi^2(m_y-1)$ are independent

and so

$$\frac{(m_x-1)S_x^2}{\sigma^2} + \frac{(m_y-1)S_y^2}{\sigma^2} \sim \chi^2(m_x + m_y - 2)$$

these two quantities are the same!

Defining the pooled variance as $S_p^2 = \dfrac{(m_x-1)S_x^2 + (m_y-1)S_y^2}{m_x + m_y - 2}$, we get $U = \dfrac{(m_x+m_y-2)S_p^2}{\sigma^2} \sim \chi^2(m_x+m_y-2)$

We then observe that

$$T = \frac{Z}{\sqrt{U/(m_x+m_y-2)}} = \frac{\dfrac{\bar{X}-\bar{Y}-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma^2}{m_x}+\frac{\sigma^2}{m_y}}}}{\sqrt{\dfrac{(m_x+m_y-2)\,S_p^2}{\sigma^2}\Big/(m_x+m_y-2)}} = \frac{\bar{X}-\bar{Y}-(\mu_x-\mu_y)}{S_p\sqrt{\frac{1}{m_x}+\frac{1}{m_y}}} \sim t(m_x+m_y-2)$$

Let $t_0 = t_{\alpha/2}(m_x+m_y-2)$ and note that from

$$P(-t_0 \leq T \leq t_0) = 1-\alpha$$

we get that

$$P\left(\bar{X}-\bar{Y}-t_0 \cdot S_p\sqrt{\frac{1}{m_x}+\frac{1}{m_y}} \leq \mu_x-\mu_y \leq \bar{X}-\bar{Y}+t_0 S_p\sqrt{\frac{1}{m_x}+\frac{1}{m_y}}\right) = 1-\alpha$$

Setting $\bar{X} = \bar{x} = 65.7$, $\bar{Y} = \bar{y} = 68.2$, $S_x = 4$, $S_y = 3$, $m_x = 12$, $m_y = 15$, observing that $\alpha = 0.02$ (because we want $1-\alpha = 0.98$) and checking on the table for the t distribution for the value $t_0 = t_{\alpha/2}(m_x+m_y-2) = t_{0.01}(25) = 2.485$, we obtain the endpoints of the 98% confidence interval for $\mu_x - \mu_y$:

$$\bar{x} - \bar{y} \pm t_0 . S_p \sqrt{\frac{1}{m_x} + \frac{1}{m_y}} = 65.7 - 68.2 \pm (2.485) . (3.4756) \sqrt{\frac{1}{12} + \frac{1}{15}},$$

where we computed S_p as $S_p = \sqrt{\dfrac{(m_x-1)S_x^2 + (m_y-1)S_y^2}{m_x+m_y-2}} = \sqrt{\dfrac{11 \cdot (4)^2 + 14 \cdot (3)^2}{25}} \approx 3.4756$

Completing the evaluations, we obtain the interval $[-5.845, 0.845]$.