Sec. 8.2 Ex 11

Let $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$, where σ_x^2 and σ_y^2 are both unknown and not necessarily equal.

a) We want to test

$$H_0 : \mu_x = \mu_y \quad \text{vs} \quad H_1 : \underbrace{\mu_x > \mu_y}_{\text{same as } \mu_x - \mu_y > 0}$$

given two samples of respective sizes $n_x = 90$ and $n_y = 110$. Since the distributions of X and Y are normal (and we will also assume independence), σ_x^2 and σ_y^2 are unknown and possibly not equal and both sample sizes are large ($n_x > 30$ and $n_y > 30$), we have that, under H_0, the test statistic is

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} \quad \text{is approximately } N(0,1)$$

We reject H_0 if the value z of the test statistic Z is such that $z \geq z_\alpha$

Picking a significance level $\alpha = 0.05$, we obtain the critical region

$$z \geq z_\alpha = z_{0.05} = 1.645$$

Since we are given that

$$\bar{x} = 8.1, \quad s_x = 0.117, \quad n_x = 90$$

and

$$\bar{y} = 8.07, \quad s_y = 0.054, \quad n_y = 110$$

we obtain

$$= \frac{8.1 - 8.07}{\sqrt{\frac{(0.117)^2}{90} + \frac{(0.054)^2}{110}}} \approx 2.245 > z_{0.05} = 1.645,$$

and we reject H_0 at significance level $\alpha = 0.05$.

b) The p-value is the probability, under H_0, of observing a value more extreme than the one yielded by the collected samples, that is:

$$\text{p-value} = P(Z \geq z \mid H_0 \text{ true})$$

$$= P(Z \geq 2.245)$$

$$= 1 - P(Z < 2.245) = 1 - \Phi(2.245)$$

$$= 1 - \frac{1}{2}(0.9875 + 0.9878)$$

$$= 1 - 0.98765$$

$$\approx 0.01235$$