Løsningsforslag (ST1201/ST6201 2016)

1.

$$X_1, X_2, ..., X_n \sim N(\mu, \sigma^2)$$

$$H_0 : \mu = \mu_0 \quad H_1 : \mu > \mu_0$$

where $\sigma = 0.05$, $\mu_0 = 7.45$.

a) The test statistic is

$$\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma}$$

which has the standard normal distribution under H_0. H_0 is rejected if

$$\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma} \geq z_\alpha.$$

In our case $n = 20$, $\bar{X} = 7.47$, $\mu_0 = 7.45$, $\sigma = 0.05$, $\alpha = 0.05$, $z_\alpha = 1.645$.

$$\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma} = 1.788.$$

H_0 is rejected.

b) We have to find

$$P\left(\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma} \geq z_\alpha\right)$$

when $X_i \sim N(\mu_1, \sigma^2)$ where $\mu_1 = 7.47$.

$$P\left(\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma} \geq z_\alpha\right) = P\left(\sqrt{n}\frac{\bar{X} - \mu_1 + \mu_1 - \mu_0}{\sigma} \geq z_\alpha\right) =$$

$$= P\left(\sqrt{n}\frac{\bar{X} - \mu_1}{\sigma} \geq z_\alpha - \sqrt{n}\frac{\mu_1 - \mu_0}{\sigma}\right) = P\left(Z \geq z_\alpha - \sqrt{n}\frac{\mu_1 - \mu_0}{\sigma}\right) =$$

$$= 1 - \Phi\left(z_\alpha - \sqrt{n}\frac{\mu_1 - \mu_0}{\sigma}\right) = \Phi\left(\sqrt{n}\frac{\mu_1 - \mu_0}{\sigma} - z_\alpha\right) =$$

$$= \Phi(0.143) = 0.5557$$

(here $\Phi(z)$ is the standard normal distribution function).

c) n must be such that

$$\Phi\left(\sqrt{n}\frac{\mu_1 - \mu_0}{\sigma} - z_\alpha\right) \geq 0.8$$

or

$$\Phi(0.4\sqrt{n} - 1.645) \geq 0.8.$$

From the table we find $0.4\sqrt{n} - 1.645 \geq 0.85$ and hence $n \geq 39$.

2.

$$X_1, X_2, ..., X_n \sim N(1, \theta)$$

$$H_0 : \theta = 1 \quad H_1 : \theta \neq 1$$

a) $\hat{\theta} \sim \chi_n^2$ under H_0. H_0 is rejected if

$$n\hat{\theta} \leq \chi^2_{1-\alpha/2,n} \text{ or } n\hat{\theta} \geq \chi^2_{\alpha/2,n}.$$

The power function is

$$\pi(\theta) = P_\theta(n\hat{\theta} \leq \chi^2_{1-\alpha/2,n}) + P_\theta(n\hat{\theta} \geq \chi^2_{\alpha/2,n}) =$$

$$= P_\theta(\frac{1}{\theta}n\hat{\theta} \leq \frac{1}{\theta}\chi^2_{1-\alpha/2,n}) + P_\theta(\frac{1}{\theta}n\hat{\theta} \geq \frac{1}{\theta}\chi^2_{\alpha/2,n}) =$$

$$= P_\theta(V \leq \frac{1}{\theta}\chi^2_{1-\alpha/2,n}) + P_\theta(V \geq \frac{1}{\theta}\chi^2_{\alpha/2,n})$$

where V has χ_n^2-distribution.

b) First we find the likelihood function $L(\theta)$ and the maximum likelihood estimator of θ.

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(X_i-1)^2}{2\theta}} = (2\pi)^{-n/2}\theta^{-n/2}e^{-\frac{1}{2\theta}\sum(X_i-1)^2}.$$

$$\ln L(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(X_i - 1)^2.$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(X_i - 1)^2 = 0.$$

The maximum likelihood estimator is

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - 1)^2.$$

The likelihood ratio is

$$\lambda = \frac{L(1)}{L(\hat{\theta})} = (\hat{\theta})^{n/2}e^{(n/2)(1-\hat{\theta})}.$$

λ is not monotone in $\hat{\theta}$ therefore these test statistics give different tests.

3.

a) The least squares estimators of α and β are such values of these parameters that minimize the sum

$$\sum_{i=1}^{n}(Y_i - \alpha - \beta x_i)^2.$$

These estimators are

$$\hat{\beta} = \frac{\sum(x_i - \bar{x})Y_i}{\sum(x_i - \bar{x})^2}$$

and

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}.$$

b) We have to test

$$H_0 : \beta = 0 \quad H_1 : \beta \neq 0.$$

The test statistic is

$$T = \frac{\hat{\beta}}{\sqrt{S^2/\sum_{i=1}^{10}(x_i - \bar{x})^2}}$$

which (under H_0) has t-distribution with $10 - 2 = 8$ degrees of freedom. H_0 is rejected if $|T| \geq t_{0.025,8}$. In our case

$$T = \frac{1.112}{\sqrt{2.3/4.1}} = 1.49.$$

$t_{0.025,8} = 2.306$. H_0 is not rejected.

4.
a)

$$f_{Y_i}(y) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(y-\alpha-\beta(x_i-\bar{x}))^2}{2\sigma_0^2}}.$$

The likelihood function is

$$L(\alpha, \beta) = \prod_{i=1}^{n} f_{Y_i}(Y_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(Y_i-\alpha-\beta(x_i-\bar{x}))^2}{2\sigma_0^2}} =$$

$$= (2\pi)^{-n/2}\sigma_0^{-n} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(Y_i-\alpha-\beta(x_i-\bar{x}))^2}.$$

The log-likelihood is

$$\ln L(\alpha, \beta) = -\frac{n}{2}\ln(2\pi) - n\ln\sigma_0 - \frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(Y_i - \alpha - \beta(x_i - \bar{x}))^2.$$

The derivatives are

$$\frac{\partial \ln L}{\partial \alpha} = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(Y_i - \alpha - \beta(x_i - \bar{x}))$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(Y_i - \alpha - \beta(x_i - \bar{x}))(x_i - \bar{x})$$

Solving the equations

$$\frac{\partial \ln L}{\partial \alpha} = 0 \quad \frac{\partial \ln L}{\partial \beta} = 0,$$

we obtain

$$\hat{\alpha} = \bar{Y},$$

$$\hat{\beta} = \frac{\sum(x_i - \bar{x})(Y_i - \bar{Y})}{\sum(x_i - \bar{x})^2} = \frac{\sum(x_i - \bar{x})Y_i}{\sum(x_i - \bar{x})^2}.$$

$$\mathrm{Var}\hat{\beta} = \mathrm{Var}\left(\frac{\sum(x_i-\bar{x})Y_i}{\sum(x_i-\bar{x})^2}\right) = \frac{1}{(\sum(x_i-\bar{x})^2)^2}\sum \mathrm{Var}((x_i-\bar{x})Y_i) =$$
$$= \sigma_0^2 \frac{\sum(x_i-\bar{x})^2}{(\sum(x_i-\bar{x})^2)^2} = \frac{\sigma_0^2}{\sum(x_i-\bar{x})^2}.$$

b)

$$\hat{\beta} \sim N\left(\beta, \frac{\sigma_0^2}{\sum(x_i-\bar{x})^2}\right).$$

Normal because it is a linear combination of independent normally distributed random variables. The variance is found above.

$$E\hat{\beta} = \frac{\sum(x_i-\bar{x})EY_i}{\sum(x_i-\bar{x})^2} = \frac{\sum(x_i-\bar{x})(\alpha+\beta(x_i-\bar{x}))}{\sum(x_i-\bar{x})^2} =$$
$$= \alpha\frac{\sum(x_i-\bar{x})}{\sum(x_i-\bar{x})^2} + \beta\frac{\sum(x_i-\bar{x})^2}{\sum(x_i-\bar{x})^2} = \alpha\cdot 0 + \beta\cdot 1 = \beta.$$

Then

$$\sqrt{\sum(x_i-\bar{x})^2}\frac{\hat{\beta}-\beta}{\sigma_0} \sim N(0,1),$$
$$P\left(-z_{\delta/2} \le \sqrt{\sum(x_i-\bar{x})^2}\frac{\hat{\beta}-\beta}{\sigma_0} \le z_{\delta/2}\right) = 1-\delta,$$

and therefore a $(1-\delta)$-confidence interval is

$$\left[\hat{\beta} - z_{\delta/2}\frac{\sigma_0}{\sqrt{\sum(x_i-\bar{x})^2}}, \hat{\beta} + z_{\delta/2}\frac{\sigma_0}{\sqrt{\sum(x_i-\bar{x})^2}}\right].$$

5.
a)

$$df(Tr.) = SSTR/MSTR = 24.8/8.16 = 3$$
$$SSE = MSE\cdot df = 5.1\cdot 40 = 204$$
$$SSTOT = SSTR + SSE = 24.48 + 204 = 228.48$$
$$df(Tot.) = df(Tr.) + df(E) = 3 + 40 = 43$$
$$F = MSTR/MSE = 8.16/5.1 = 1.6$$

Thus the filled ANOVA table is

Source	df	SS	MS	F
Treatment	3	24.48	8.16	1.6
Error	40	204	5.1	
Total	43	228.48		

$F_{0.95,3,40} = 2.84 > 1.6$ (observed F). Therefore H_0 is not rejected.