Løsningsforslag (ST1201/ST6201 2019)

1. It easy to see that

$$E\hat{\mu}_1 = E\hat{\mu}_2 = \mu$$

i.e. the both estimators are unbiased. Find the variances

$$Var(\mu_1) = (2 \cdot \frac{1}{9} + \frac{1}{9})\sigma^2 = \frac{\sigma^2}{3},$$

$$Var(\mu_2) = (2 \cdot \frac{1}{4} + \frac{1}{16})\sigma^2 = \frac{9\sigma^2}{16}.$$

The first estimator has less variance, i.e. it is more efficient and therefore more preferable.

2.

$$X_1, X_2, ..., X_n \sim N(0, \sigma^2)$$
$$H_0 : \sigma^2 = 1 \quad H_1 : \sigma^2 \neq 1$$

a) $nT \sim \chi^2_n$ under H_0. H_0 is rejected if

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

The power function is

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

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

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

where V has χ^2_n-distribution.

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

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

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

$$\frac{\partial \ln L(\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n} X_i^2 = 0.$$

The maximum likelihood estimator is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = T.$$

The likelihood ratio is

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

λ is not one-to-one with T therefore these test statistics give different tests.

3.
a) The log-likelihood function is

$$\ln L(\beta) = -\frac{n}{2}\ln(2\pi) - n\ln\sigma_0 + \ln\left(\prod_{i=1}^{n}\frac{1}{x_i}\right) - \frac{1}{\sigma_0^2}\sum_{i=1}^{n}\frac{(Y_i - \beta x_i)^2}{x_i^2}.$$

Its derivative is

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

Solution of the equation

$$\sum_{i=1}^{n}\frac{Y_i - \beta x_i}{x_i} = 0$$

is MLE. This solution is

$$\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i}{x_i}.$$

$$E\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n}\frac{EY_i}{x_i} = \frac{1}{n}\sum_{i=1}^{n}\frac{\beta x_i}{x_i} = \beta.$$

$$Var\hat{\beta} = \frac{1}{n^2}\sum_{i=1}^{n}\frac{VarY_i}{x_i^2} = \frac{\sigma_0^2}{n}.$$

b) $\hat{\beta}$ has a normal distribution because it is a linear combination of independent random variables having normal distributions. So

$$\hat{\beta} \sim N(\beta, \sigma_0^2/n).$$

Therefore

$$\sqrt{n}\frac{\hat{\beta} - \beta}{\sigma_0} \sim N(0,1).$$

Then

$$P\left(-z_{\alpha/2} \le \sqrt{n}\frac{\hat{\beta} - \beta}{\sigma_0} \le z_{\alpha/2}\right) = 1 - \alpha.$$

The $(1-\alpha)-$confidence interval is

$$\left[\hat{\beta} - z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}, \hat{\beta} + z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right]$$

4.

La $X = X_1 + X_2$. Da $X \sim \chi^2_{2n}$. Fra "Tabeller og formler ..." er sannsynlighetstettheten (ST) til en χ^2-fordelt stokastisk variabel X med $2n$ frihetsgrader (n=1,2,...) gitt som

$$f_X(x) = c_n x^{n-1} e^{-x/2}$$

for $x \geq 0$ og 0 ellers, hvor $c_n = 1/(2^n(n-1)!)$. Siden $f_X(x)$ er en ST, må det gjelde at

$$c_n \int_0^\infty x^{n-1} e^{-x/2} dx = 1, \quad n = 1, 2, ...$$

Dermed er

$$\mathrm{E}(X^{-1}) = c_n \int_0^\infty x^{-1} x^{n-1} e^{-x/2} dx = c_n \int_0^\infty x^{(n-1)-1} e^{-x/2} dx =$$

$$= \frac{c_n}{c_{n-1}} c_{n-1} \int_0^\infty x^{(n-1)-1} e^{-x/2} dx = \frac{c_n}{c_{n-1}} = \frac{1}{2(n-1)}, \quad n = 2, 3, ...$$

Tisvarende,

$$\mathrm{E}(X^{-2}) = c_n \int_0^\infty x^{-2} x^{n-1} e^{-x/2} dx = c_n \int_0^\infty x^{(n-2)-1} e^{-x/2} dx =$$

$$= \frac{c_n}{c_{n-2}} c_{n-2} \int_0^\infty x^{(n-2)-1} e^{-x/2} dx = \frac{c_n}{c_{n-2}} = \frac{1}{4(n-1)(n-2)}, \quad n = 3, 4, ...$$

5.
a)

$$df(Tr.) = SSTR/MSTR = 18.1/3.62 = 5$$

$$SSE = SSTOT - SSTR = 247.7 - 18.1 = 229.6$$

$$df(Tot.) = df(Tr.) + df(E) = 5 + 80 = 85$$

$$MSE = SSE/df(Err.) = 229.6/80 = 2.87$$

$$F = MSTR/MSE = 3.62/2.87 = 1.26$$

Thus the filled ANOVA table is

Source	df	SS	MS	F
Treatment	5	18.1	3.62	1.26
Error	80	229.6	2.87	
Total	85	247.7		

The number of observations is 86, the number of samples is 6. They (the samples) cannot have all the same size.

b) The observed value of the test statistic is 1.26. The critical value is $f_{0.05,5,80} = 2.33$. The null hypothesis is not rejected.

6.
a) The problem

$$H_0 : \mu_1 = \mu_2, \ H_1 : \mu_1 \neq \mu_2.$$

Since the variances are unknown but equal, the test statistic is

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{1/n_1 + 1/n_2}},$$

where S_p^2 is the pooled empirical variance. Under H_0 $T \sim t_{n_1+n_2-2}$. In our case the observed value of the test statistic is $t = -0.471$. Since $t_{0.025,13} = 2.16$, H_0 is not rejected.

b) For the given data the corresponding ranks are in the table.

3	5	1	7	11.5	14	10	
6	8	2	4	9	15	11.5	13

The test statistic is

$$Z = \frac{W - n_1(n_1+n_2+1)/2}{\sqrt{n_1 n_2 (n_1+n_2+1)/12}},$$

where

$$W = \sum_{i=1}^{n_1} R_i$$

(R_i is the rank of the $i-$th observation). Under H_0 the test statistic has the standard normal distribution.

In our case

$$z = \frac{51.5 - 56}{8.64} = -0.52.$$

Since $z_{0.025} = 1.96$, H_0 is not rejected.