Løsningsforslag (ST1201/ST6201 høst 2014)

1.

a) Let X be the number of hits. X has the binomial distribution with parameters (n, p) where $n = 20$. The following hypothesis is tested

$$H_0 : p = p_0 = 0.6,\ H_1 : p > p_0 = 0.6.$$

Under H_0 the test statistic

$$Z = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}}$$

has the standard normal distribution (approximately). H_0 is rejected if $Z \geq z_\alpha$. In our case $z_\alpha = 1.645$ and the observed value of Z is 2.7. H_0 is rejected.

b) The $p-$value is equal to $P(X \geq 18)$ where X has the binomial distribution with parameters (20,0.6). This probability can be calculated either directly (exactly) or using the normal approximation.

1. Directly

$$P(X \geq 18) = \frac{20!}{18!2!}0.6^{18}0.4^2 + \frac{20!}{19!1!}0.6^{19}0.4^1 + \frac{20!}{20!0!}0.6^{20}0.4^0 = 0.0035.$$

2. Using normal approximation

$$P(X \geq 18) = P\left(Z \geq \frac{18 - 20 \cdot 0.6}{\sqrt{20 \cdot 0.6 \cdot 0.4}}\right) =$$

$$= P(Z \geq 2.7) = P(Z \leq -2.7) = 0.0035.$$

2.

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

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

b)

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

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

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

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

Therefore

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

Then

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

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

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

3.

a) The randomized block design should be used. The model is

$$Y_{ij} = \mu_j + \beta_i + \epsilon_{ij},\ \epsilon_{ij} \sim N(0,\sigma^2),$$

where μ_j is the treatment effect of instrument number j, β_i is the block effect of object number i, ϵ_{ij} are random errors.

We have $k-1=4$, $b-1=5$, $(b-1)(k-1)=20$, $n-1=29$. These are elements of the "df"- column. The first element of the "F"- column is such a value f that $P(F_{4,20} > f) = 0.025$, where $F_{m,n}$ is a random variable having the $F-$distribution with m and n degrees of freedom. We find from the table of quantiles of $F-$distributions that it is 3.51. The second element of the "F"- column is such a value f that $P(F_{5,20} > f) = 0.05$. It is 2.71. Since MS = SS/df and F = MSTR/MSE (for treatments) F = MSB/MSE (for blocks), and SSTOT = SSTR + SSB + SSE, we finally obtain

Kilde	df	SS	MS	F	P-verdy
Instrument	4	8	2	3.51	0.025
Objekt	5	7.7	1.54	2.71	0.05
Error	20	11.4	0.57		
Total	29	27.1			

b) H_0 for the $p-$value in the first line is

$$H_0: \ \mu_1 = \mu_2 = ... = \mu_5.$$

H_0 for the $p-$value in the second line is

$$H_0: \ \beta_1 = \beta_2 = ... = \beta_6.$$

The institute is interested in the first $p-$value (differences in instruments). Since the $p-$value is smaller than the significance level, the null hypothesis is rejected.

4.

a) The hypothesis

$$H_0 : \mu_1 = \mu_2$$

is tested versus the alternative

$$H_1 : \mu_1 > \mu_2.$$

We can use the pared $t-$test. The test statistic

$$T = \sqrt{n}\frac{\bar{D}}{S_D}.$$

The observed value of the test statistic is

$$t_{\text{obs}} = \sqrt{15}\frac{21.53}{38.29} = 2.177.$$

This is a one-sided test, the null hypothesis is rejected if $t_{\text{obs}} \geq t_{\alpha,n-1}$. In our case

$$t_{\text{obs}} = 2.177 > 1.761 = t_{0.05,14}.$$

The null hypothesis is rejected.

b) We use the sign test. Let D be the number of differences $D_i = X_{1i} - X_{2i}$ which are positive. The null hypothesis H_0 is rejected if the test statistic

$$\frac{D - n/2}{\sqrt{n/4}},$$

which under H_0 has (approximately) the standard normal distribution, is greater than z_α. In our case, $z_\alpha = 1.645$, and the observed value of the test statistic is

$$\frac{13 - 7.5}{\sqrt{15/4}} = 2.84.$$

H_0 is rejected.

5. Denote probabilities

$$p_{11} = P(\text{good husband, good wife}),$$

$$p_{12} = P(\text{good husband, bad wife}),$$

$$p_{21} = P(\text{bad husband}, \text{ good wife}),$$
$$p_{22} = P(\text{bad husband}, \text{ bad wife}),$$
$$p_1 = P(\text{good husband}),$$
$$p_2 = P(\text{bad husband}),$$
$$q_1 = P(\text{good wife}),$$
$$q_2 = P(\text{bad wife}).$$

(It is clear that $p_1 + p_2 = 1$, $q_1 + q_2 = 1$, $p_{11} + p_{11} + p_{21} + p_{22} = 1$.) It is tested the null hypothesis $H_0 : p_{ij} = p_i q_j, i = 1, 2, j = 1, 2$. We are given the contingency table

X_{11}	X_{12}
X_{21}	X_{22}

The test statistic (we use notations of the textbook) is

$$D_2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(X_{ij} - n\hat{p}_i \hat{q}_j)^2}{n\hat{p}_i \hat{q}_j}$$

where $\hat{p}_i, \hat{q}_j$ are estimators (MLE) of p_i, q_j. H_0 is rejected if

$$D_2 \geq \chi^2_{1-\alpha,1}.$$

In our case $D_2 = 1.02$, $\chi^2_{1-\alpha,1} = 3.841$. The null hypothesis is not rejected (that means that there is no reason to believe that the temperament of husband and wife are dependent of each other).