**Solutions. TMA4255, 2019 (June)**

**1.**

**a)** The model is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},\ i = 1, 2, ..., a,\ j = 1, 2, ..., b,\ k = 1, 2, ..., n$$

($a = 3,\ b = 4,\ n = 2$), where $\alpha_i$ is effect of factor A, $\beta_j$ is effect of factor B, $(\alpha\beta)_{ij}$ is effect of interaction,

$$\sum_{i=1}^{a} \alpha_i = 0,\ \sum_{j=1}^{b} \beta_j = 0,\ \sum_{i=1}^{a} (\alpha\beta)_{ij} = 0,\ \sum_{j=1}^{b} (\alpha\beta)_{ij} = 0,$$

$\epsilon$-s are independent and normally distributed with zero expectation and some (unknown) variance $\sigma^2$.

Test on the interaction:

$$H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = ... = (\alpha\beta)_{ab} = 0$$

$$H_1 : (\alpha\beta)_{ij} \neq 0 \text{ for at least one } (i, j).$$

Test statistic is

$$F = \frac{MS(AB)}{MSE} = \frac{SS(AB)/(a-1)(b-1)}{MSE/ab(n-1)}.$$

$H_0$ is rejected for large values of $F$. Under $H_0$ the test statistic has $F$-distribution with $(a-1)(b-1)$ and $ab(n-1)$ degrees of freedom. $H_0$ is rejected if $F > f_{\alpha,(a-1)(b-1),ab(n-1)}$. In our case $F = 2.79$, $f_{0.05,6,12} = 3.00$. $H_0$ is not rejected (there is no interaction between the two factors).

The main effects are significant because

$$F = 93.52 > 3.89 = f_{0.05,2,12}$$

for factor A,

$$F = 1330.48 > 3.49 = f_{0.05,3,12}$$

for factor B.

**b)**

$$SSE = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (Y_{ijk} - \bar{Y}_{ij\cdot})^2,$$

$$SS(AB) = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot j\cdot} + \bar{Y}_{\cdot\cdot\cdot})^2,$$

$$S^2 = \frac{SSE}{ab(n-1)} = MSE = 0.537.$$

A new estimator can be derived as follows. Since

$$\frac{SSE}{\sigma^2} \sim \chi^2_{ab(n-1)}, \quad \frac{SS(AB)}{\sigma^2} \sim \chi^2_{(a-1)(b-1)},$$

we get

$$\frac{SSE}{\sigma^2} + \frac{SS(AB)}{\sigma^2} \sim \chi^2_{(a-1)(b-1)+ab(n-1)},$$

therefore a reasonable estimator (unbiased) is

$$\hat{\sigma}^2 = \frac{SSE + SS(AB)}{(a-1)(b-1) + ab(n-1)} = \frac{SSE + SS(AB)}{18}.$$

Computed

$$\hat{\sigma}^2 = \frac{8.99 + 6.45}{18} = 0.8578.$$

**2.**

**a)** The observations $Y_{ijk}$ can be renumbered from 1 to 24, and for each of them, the corresponding $x_{1i}$ and $x_{2j}$ are values of the two explanatory variables.

Assumptions: the variables $\epsilon_{ijk}$ are independent and normally distributed with zero mean and some (unknown) variance $\sigma^2$.

Variation explained by the model:

$$R^2 = \frac{SSR}{SST} = \frac{2242.6}{2261.4} = 0.992.$$

Significance is tested by

$$F = \frac{SSR/2}{SSE/21} = 1251.4$$

This is significant by any reasonable significance level.

For testing

$$H_0 : \beta_r = 0 \text{ versus } H_1 : \beta_r \neq 0, \;\; r = 1, 2,$$

test statistics

$$T_r = \frac{\hat{\beta}_r}{SD(\hat{\beta}_r)}$$

are used. Under $H_0$ each of these statistics has $t$-distribution with 21 degrees of freedom, therefore $H_0$ is rejected if $|T_r| > t_{0.005,21} = 2.831$. So, both hypotheses are rejected.

The tests are essentially about the main effects, so the results of Problem 1 are confirmed.

**3.**

Generally $P(H_0 \text{ is rejected}) = P(\bar{X} > c)$.

**a)**

$$\alpha = P(\bar{X} > c | \mu = 0) = P\left(\sqrt{n}\frac{\bar{X}}{\sigma} > \sqrt{n}\frac{c}{\sigma} | \mu = 0\right) = P\left(Z > \sqrt{n}\frac{c}{\sigma}\right),$$

therefore

$$\sqrt{n}\frac{c}{\sigma} = z_\alpha,$$

and

$$c = \frac{\sigma}{\sqrt{n}} z_\alpha.$$

We have

$$0.5 = P\left(\bar{X} > \frac{\sigma}{\sqrt{n}} z_\alpha | \mu = 1\right) = P\left(\sqrt{n}\frac{\bar{X}-1}{\sigma} > \sqrt{n}\frac{\sigma z_\alpha/\sqrt{n}-1}{\sigma} | \mu = 1\right) =$$

$$= P(Z > z_\alpha - \sqrt{n}/\sigma),$$

hence $z_\alpha - \sqrt{n}/\sigma = z_{0.5} = 0$, i.e. $z_\alpha = 2$, and $\alpha = 0.0228$.

**b)** We have (see solution of (a))

$$P\left(\bar{X} > \frac{\sigma}{\sqrt{n}} z_\alpha | \mu = 2\right) = P(Z > z_\alpha - 2\sqrt{n}/\sigma) = P(Z > -2) = P(Z < 2) = \Phi(2) = 0.9772.$$

**4.**

**a)** Test $H_0$ : independence vs. $H_1$ : not independence. Test statistic

$$\chi^2 = \sum_{j=1}^{3}\sum_{i=1}^{2} \frac{(o_{ij} - e_{ij})^2}{e_{ij}},$$

where $o_{ij}$ and $e_{ij}$ are observed and expected frequencies, respectively. $H_0$ is rejected if $\chi^2 > \chi^2_{0.01,2}$.

Find expected frequencies:

| | Non-smoker | Moderate smoker | Heavy smoker | |
|---|---|---|---|---|
| High blood pressure | 33.24 | 30.31 | 24.44 | 88 |
| Normal blood pressure | 34.76 | 31.69 | 25.56 | 92 |
| | 68 | 62 | 50 | 180 |

Thus, $\chi^2 = 16.982$, $\chi^2_{0.01,2} = 9.21$. $H_0$ is rejected.

**5.**

**a)** $P$-value is wrong. It must be between 0 and 1 (propability!). The right value is $p = P(|Z| > 0.31) = 2P(Z < -0.31) = 2 \cdot 0.3783 = 0.7566$ (0,760 in the MINITAB output).

**b)** If the confidence level is $1-\alpha$, then the $100(1-\alpha)\%$ confidence interval is

$$\left[\bar{Y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{Y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

Since 95% interval is [-0.234,0.320] and $z_{0.025} = 1.96$, $z_{0.00005} = 3.891$, the 99.99% interval is [-0.507,0.593].