ANSWERS

TAMS38 - Experimental Design and Biostatistics, 4 p / 6 ph
Examination on 15 August 2017 kl 14-18.

1 **a)** Model: $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$, where $\epsilon_{ij} \sim N(o, \sigma)$, $j = 1, ..., n_i$, $i = 1, ...4$.

$\bar{y}_{..} = 3.72$

$SS_{TR} = \sum_{i=1}^{4} n_i(\bar{y_{i.}} - \bar{y}..)^2 \approx 0.44$, where $n_1 = 4, n_2 = 5, n_3 = 3, n_4 = 5$ and $\bar{y_{1.}} = 3.66, \bar{y_{2.}} = 3.628, \bar{y_{3.}} = 3.56$ and $\bar{y_{4.}} = 3.966$.

$\hat{\sigma^2} = \frac{SSE}{N-a} \approx 0.06$, where $SSE = \sum_{i=1}^{4}(n_i - 1)s_i^2 \approx 0.77$, $N = 17$ and $a = 4$.

b) One way ANOVA

Source	DF	SS	F
Diets	3	0.44	2.48
Error	13	0.77	
Total	16		

$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ against H_1 : at least one $\tau_i \neq 0$

Test statistic

$$\upsilon = \frac{SS_{Diets}/3}{SS_E/13} \approx 2.48$$

$\upsilon < c = F(3, 13) = 3.415$, so we don't reject H_0, i.e. 4 different diets don't effect liver weights.

2 Kruskal - wallis test

H_0 : There is no significant difference between 4 diets.
H_1 : There is significant difference between 4 diets.

	Treatment		rank r_{ij}
1	2	3	4
3.52 $r_.4.5$	3.47 $r_.3$	3.54 $r_.6$	3.74 $r_.11$
3.36 $r_.1$	3.73 $r_.10$	3.52 $r_.4.5$	3.83 $r_.12$
3.57 $r_.7$	3.38 $r_.2$	3.61 $r_.8$	3.87 $r_.13.5$
4.19 $r_.16$	3.87 $r_.13.5$		4.08 $r_.15$
	3.69 $r_.9$		4.31 $r_.17$
$s_1 = 28.5$	$s_1 = 37.5$	$s_1 = 18.5$	$s_1 = 68.5$

$$T = T_{\text{ties}} \frac{(N-1)(s_a - C)}{s_r - C} \approx 6.28,$$

where $s_a = \sum_{i=1}^{a} \frac{s_i^2}{n_i} \approx 1536.85$, $C = \frac{N(N+1)}{4} = 1377$ and $s_r = \sum_i \sum_j r_{ij}^2 = 1784$.
$T = 6.28 < \chi^2_{0.01}(3) = 11.32$, we don't reject H_0.

3 **a)** Response variables, C1, C2, C3. Factors: week 1,2,3. Machine: 1,2. Operator: A,B.

b) i) NO ii) NO iii) YES iv) YES v) NO

c) Tukey interval gives

$$I_{\mu_i-\mu_j} = (\bar{y}_{i..} - \bar{y}_{j..}) \pm q_{0.05}(a, dfE)\frac{s}{\sqrt{n}} = (\bar{y}_{i..} - \bar{y}_{j..}) \pm 3.53\frac{0.89}{\sqrt{10}} = (\bar{y}_{i..} - \bar{y}_{j..}) \pm 0.993.$$

We can say that there is a significant difference between week 1 and week 3 since $\mu_1 < \mu 3$.

d) t- interval gives

$$I_\mu = \bar{y}_{...} \pm t_{0.01}(dfE)\frac{s}{\sqrt{30}} = 5.507 \pm 2.49\frac{0.89}{\sqrt{30}} \approx (5.10, 5.91).$$

4 **a)** nr.3 B nr.5 C

b) If we choose $E = ABCD$, then

$$E = 1 : (1), ab, ac, bc, ad, bd, cd, abcd.$$
$$E = -1 : a, b, c, abc, d, abd, acd, bcd.$$

c) Model:

$$Y = \mu + \tau_i + \beta_j + \gamma_k + \delta_l + (\beta\gamma)_{jk} + \epsilon_{ijkl},$$

where $\epsilon_{ijkl} \sim N(o, \sigma)$, $\sum \tau_i = 0, \sum \beta_i = 0, \sum \gamma_k = 0, \sum \delta_l = 0$ and $\sum_j (\beta\gamma)_{jk} = 0, \sum_k (\beta\gamma)_{jk} = 0$.

$$R^2 = \frac{SS_{TOT} - SS_E}{SS_{TOT}} = 93.4\%.$$

d) Tukey interval gives

$$I_{\tau_1-\tau_{-1}} = (44.538 - 33.425) \pm q_{0.05}(2, 10)\frac{s}{\sqrt{8}} = (44.538 - 33.425) \pm 3.15\frac{s}{\sqrt{8}} = 11.113 \pm 9.46,$$

and

$$I_{(\mu)_{jk}-(\mu)_{lm}} = (\bar{y}_{.jk.} - \bar{y}_{.lm.}) \pm q_{0.05}(4, 10)\frac{s}{\sqrt{4}} = (\bar{y}_{.jk.} - \bar{y}_{.lm.}) \pm 4.33\frac{s}{\sqrt{4}} = (\bar{y}_{.jk.} - \bar{y}_{.lm.}) \pm 18.40.$$

The best combination of factors is $A = 1, B = -1, C = 1$.