Lösningsförslag

Uppgift 1

3) t = är den totala skadan i tkr

$\hat{t}_{yr} = N\hat{B}\bar{x}_u$

$\hat{B} = \frac{\sum y}{\sum x} = \frac{612}{2.71} = 225.83$

$N = 100 \quad \bar{x}_u = 0.62$

så $\hat{t}_{yr} = 14001.46$

$SE[\hat{t}_{yr}]^2 = 100^2 \cdot \frac{s_e^2}{5}\left(1 - \frac{5}{100}\right)$

$s_e^2 = \frac{1}{4}\sum(y_i - \hat{B}x_i)^2 =$

$= \frac{1}{4}\left(\sum y_i^2 + \hat{B}^2 \sum x_i^2 - 2\hat{B}\sum x_i y_i\right)$

$= \frac{1}{4}(76520 + 225.83^2 \cdot 1.6923$

$- 2 \cdot 225.83 \cdot 350.48) = 1132\ 032\ 644$

$SE[\hat{t}_{yr}] = 1466.5817$

95% KI för t

14001 ± 2874.5

Uppgift 2

a) Andelen *Game of Thrones* -tittare i delgruppen kvinnor, P_k skattas med $\hat{P}_k = \frac{240}{1200} = 0{,}2$, varför konfidensintervallet blir (om fcp ignoreras):

$$N_k * \hat{P}_k \pm 1.96 * N_k\sqrt{\frac{\hat{P}_k(1-\hat{P}_k)}{n_k}} = 2\,000\,000 * 0{,}2 \pm 3\,920\,000 * \sqrt{\frac{0{,}2 * 0{,}8}{1200}}$$
$$= \mathbf{400\,000 \pm 45264{,}3}$$

b) (Se sid 134-135 i Lohr, andra upplagan). Utan kännedom om antalet tv-serietittande kvinnor skattas antalet kvinnliga tittare av serien med:

$$\hat{t}_k = N * \bar{u} = 4\,000\,000 * \frac{240}{2000} = 480\,000$$
$$där\ u_i = 1\ om\ kvinnlig\ tittare\ på\ Game\ of\ Thrones.\ 0\ annars$$

Medelfelet blir:

$$SE(\hat{t}_k) = N\sqrt{\left(1-\frac{n}{N}\right)\frac{s_u^2}{n}} =$$
$$= 4\,000\,000 * \sqrt{\left(1-\frac{2000}{4000000}\right)\frac{0{,}105653}{2000}} = \mathbf{29065{,}5}$$

där:

$$s_u^2 = \frac{1}{n-1}\left(\sum u_i^2 - \frac{(\sum u_i)^2}{n}\right) =$$
$$= \frac{1}{1999}\left(240 - \frac{240^2}{2000}\right) = 0{,}105653$$

Konfidensintervallet blir då:

$$\mathbf{480\,000 \pm 56968{,}4} = 1{,}96 * 29065{,}5$$

Uppgift 3

4. Beteckna den totala avkastningen med t_y.

a) Skattningen som utnyttjar bara y-värdena är
$\hat{t}_y = N \cdot \overline{y} = 80 \cdot 158.9 = 12712$.
Skattade variansen är $N^2(1-\frac{n}{N})\frac{s_y^2}{n} = 80^2(1-\frac{20}{80})\frac{56^2}{20} = 752640$, varför variationskoefficienten blir $CV = \frac{\sqrt{752640}}{12712} = 6.8\,\%$

b) Regressionsskattningen är
$\hat{t}_{yreg} = N(\hat{B}_0 + \hat{B}_1 \cdot \overline{x}_U) = 80(68.365 + 36.194 \cdot \frac{180}{80}) = 11984$
och har variansen $N^2(1-\frac{n}{N})\frac{s_e^2}{n} = 80^2(1-\frac{20}{80})\frac{5823/19}{20} = 73553.68$, varför vi får $CV = \frac{271.21}{11984} = 2.3\,\%$.
CV är nu tre gånger mindre än i a), vilket förstås beror på att avkastningssiffrorna har starkt linjärt samband med gödselsiffrorna.

c) Kvotskattning inte lämplig, eftersom sambandet inte går genom origo som vi ser (och förstår, eftersom avkastning erhålls även utan gödsling). Om man räknar på CV i detta fall så blir den t.o.m. större än i a).

Uppgift 4

4) $\hat{\bar{y}}_{reg} = \hat{B}_0 + \hat{B}_1 \bar{x}_u$

$= 39.5 + 14.5 \cdot \frac{4}{20} = \underline{42.4}$

$\bar{y} = 39.5, \quad \bar{x} = 0.5$

$\sum x_i y_i = 108 \qquad \sum x_i^2 = 2$

$e = \hat{y} - y$

$e_1 = 39.5 + 14.5 - 56 = -2$

$e_2 = 39.5 + 14.5 - 52 = 2$

$e_3 = 39.5 - 35 = 4.5$

$e_4 = 39.5 - 44 = -4.5 \qquad S_e^2 = 16.1667$

$42.4 \pm 1.96 \cdot \underbrace{\sqrt{\left(1 - \frac{4}{20}\right) \frac{16.1667}{4}}}_{1.7981}$

95% KI

42.4 ± 3.5 kg