Prüfungsbeispiel 03

3a) Ein Autofahrer testete den Treibstoffverbrauch seines Fahrzeuges: Man stelle ein lineares Regressionsmodell auf (abhängige Variable: gefahrene km, unabhängige Variable: dafür benötigte Liter Treibstoff) und schätze die Regressionsparameter.

	1	2	3	4	5	6	7	8	9	
gefahrene km	829	832	840	762	755	571	713	677	656	F(x)
Liter	65	65	61	61	56	47	59	59	53	x

a) Wählen Sie einen linearen Regressionsansatz und schätzen Sie die Parameter a, b und σ^2.

<u>Vorbereitungen:</u>

$$\sum x_i = 526$$
$$\sum x_i^2 = 31008$$
$$\bar{x} = 58{,}444$$
$$\sum y_i = 6635$$
$$\sum y_i^2 = 4958809$$
$$\bar{y} = 737{,}222$$
$$\sum x_i y_i = 391582$$
$$s_x^2 = \frac{1}{n-1}\left(\sum x_i^2 - n \cdot \bar{x}^2\right) = \frac{1}{8}(31008 - 9 \cdot 58{,}4444^2) = 33{,}277$$
$$s_y^2 = \frac{1}{n-1}\left(\sum y_i^2 - n \cdot \bar{y}^2\right) = \frac{1}{8}(4958809 - 9 \cdot 737{,}2222^2) = 8417{,}481$$

(i) Gleichung der Regressionsgeraden:

$$\hat{y}(x) = \hat{a} + \hat{b}(x - \bar{x})$$

$$\hat{a} = \bar{y} = 737{,}222$$

$$\hat{b} = \frac{s_{xy}}{s_x^2} = \frac{\dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}}{\dfrac{\sum (x_i - \bar{x})^2}{n-1}} = \frac{\dfrac{\sum x_i y_i - n\overline{xy}}{n-1}}{\dfrac{\sum x_i^2 - n \cdot \bar{x}^2}{n-1}} = \frac{\sum x_i y_i - n \cdot \overline{xy}}{\sum x_i^2 - n \cdot \bar{x}^2} = \frac{391582 - 9 \cdot 58{,}444 \cdot 737{,}222}{31008 - 9 \cdot 58{,}444^2} = \underline{\underline{14{,}285}}$$

$$\hat{y}(x) = 737{,}222 + 14{,}285(x - \bar{x})$$

(ii) Schätzung für σ^2 : Fehlervarianz

$$s^2 = \frac{n-1}{n-2}(s_y^2 - \hat{b}^2 \cdot s_x^2)$$

$$n = 9$$

$$s^2 = \frac{8}{7}(8417{,}481 - 14{,}285^2 \cdot 33{,}278)$$

$$s^2 = 1859{,}122$$

3b) Wieviele km kann der Fahrer mit seinem Auto mit einem vollen Tank (89 Liter) fahren?
Man gebe ein 97,5% - Toleranzintervall für mögliche km- Werte bei 89 Liter an.

(i) x = 89 Liter

$$\hat{y}(x) = \hat{a} + \hat{b}(x - \overline{x})$$

$$\hat{y}(89) = 737{,}222 + 14{,}285 \cdot (89 - 58{,}444)$$

$$\underline{\underline{\hat{y}(89) = 1173{,}71 \quad km}}$$

(ii) Konfidenzintervall 89 Liter Treibstoff:

$$\hat{y}_x \pm t_{n-2;1-\frac{\alpha}{2}} \cdot s \cdot \sqrt{\frac{1}{n} + \frac{(x-\overline{x})^2}{(n-1) \cdot s_x^2}}$$

$$\hat{y}_{89} \pm 2{,}365 \cdot \sqrt{1859{,}122} \cdot \sqrt{\frac{1}{9} + \frac{(89-58{,}444)^2}{8 \cdot 33{,}277}} =$$

$$\hat{y}_{89} \pm 2{,}365 \cdot 43{,}117 \cdot \sqrt{3{,}618} =$$

$$\hat{y}_{89} \pm 193{,}96 = 1173{,}71 \pm 193{,}96 \Rightarrow \underline{\underline{979{,}75 \leq \hat{y}_{89} \leq 1367{,}67}}$$

Überprüfen Sie außerdem mit einem Wahscheinlichkeitsnetz, ob die Residuen annähernd normalverteilt sind, und schätzen Sie die Parameter grafisch.

***

3c) Testen Sie, ob die gefahrenen km vom Treibstoffverbrauch abhängig sind
(Signifikanzniveau = 0.025).

Siehe Skriptum Seite 106:

$H_0 : b = 0$	$\lvert T \rvert < t_{n-2,\alpha/2}$	y hängt nicht von x ab
$H_1 : b \neq 0$	$\lvert T \rvert \geq t_{n-2,\alpha/2}$	y hängt von x ab

Teststatistik: $T = \dfrac{\hat{b} \cdot s_x \cdot \sqrt{n-1}}{s}$

$$T = 14{,}285 \cdot \sqrt{\frac{33{,}277 \cdot 8}{1859{,}122}} = 5{,}405$$

$$T = 5{,}405 > t_{7,025} = 2{,}365$$

H_0 wird verworfen $\Rightarrow$ y hängt von x ab.