

10. Lösung weitere Übungsaufgaben Statistik II WiSe 2019/2020

1. **Aufgabe:** Für 25 der größten Flughäfen wurde die Anzahl der abgefertigten Passagiere in den Jahren 2009 und 2012 erfasst. Aus den Daten (Anzahl der Passagiere in Millionen) erhält man das folgende Ergebnis.

Simple Regression - Passagiere 2012 vs. Passagiere 2009

Dependent variable: Passagiere 2012

Independent variable: Passagiere 2009

Linear model: $Y = a + b \cdot X$

Coefficients

	<i>Least Squares</i>	<i>Standard</i>	<i>T</i>	
<i>Parameter</i>	<i>Estimate</i>	<i>Error</i>	<i>Statistic</i>	<i>P-Value</i>
Intercept	9,58842	4,30419	2,22769	0,0355
Slope	0,944436	0,0864419	10,9257	0,0000

Analysis of Variance

<i>Source</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
Model	3903,43	1	3903,43	119,37	0,0000
Residual	784,802	24	32,7001		
Total (Corr.)	4688,23	25			

Predicted Values

		<i>95,00%</i>		<i>95,00%</i>	
	<i>Predicted</i>	<i>Prediction</i>	<i>Limits</i>	<i>Confidence</i>	<i>Limits</i>
<i>X</i>	<i>Y</i>	<i>Lower</i>	<i>Upper</i>	<i>Lower</i>	<i>Upper</i>
30,0		25,4698	50,3733	33,9524	41,8906
80,0	85,1433	71,8357	98,451	78,995	91,2917

- a) Wie lautet die Modellgleichung?

Lösung:

$$Y = a + b \cdot X + \varepsilon$$

Y - Passagiere 2012, X - Passagiere 2009, ε - zufällige Fehler.

- b) Wie lautet die geschätzte Regressionsfunktion?

Lösung:

$$\hat{y}(x) = \hat{a} + \hat{b} \cdot x = 9,58842 + 0,944436 \cdot x$$

c) Schätzen Sie die Varianz des Fehlers.

Lösung: Df (Total): $n - 1 = 25 \implies n = 26$ und $r = 2$ (Parameter)

$$\hat{\sigma}^2 = \frac{1}{(n - r)} SSR = \frac{1}{(26 - 2)} 784,802 = \underline{32,7}$$

(Am Anfang der Aufgabe muss es richtig heißen: „Für 26 der größten Flughäfen...“. Wäre $n = 25$, dann erhält man mit $r = 2$ (Parameter)

$$\hat{\sigma}^2 = \frac{1}{(n - r)} SSR = \frac{1}{(25 - 2)} 784,802 = \underline{34,1}$$

d) Welche Hypothesen werden in der Tabelle Coefficients getestet und wie lauten die Testentscheidungen bei $\alpha = 0,05$?

$$\begin{array}{ll} H_0 : a = 0 \text{ gegen } H_A : a \neq 0 & p = 0,0355 < \alpha \implies H_0 \text{ wird abgelehnt.} \\ H_0 : b = 0 \text{ gegen } H_A : b \neq 0 & p = 0,0000 < \alpha \implies H_0 \text{ wird abgelehnt.} \end{array}$$

Beide Parameter, a und b, sind signifikant von 0 verschieden.

e) Die Passagierzahl im Jahr 2009 war 30 Millionen. Welche Passagierzahl wird nach diesem Modell für das Jahr 2012 prognostiziert?

Lösung:

$$\hat{y}(30) = 9,58842 + 0,944436 \cdot 30 = 37,9215$$

$37,9215 \cdot 10^6$ Passagiere werden für das Jahr 2012 prognostiziert.

f) Die Passagierzahl im Jahr 2009 war 80 Millionen. Wie lautet das 95 % Prognoseintervall für die Passagierzahl im Jahr 2012?

Lösung: Prediction Limits

$$[71,8357 ; 98,451]$$