Klausurvorbereitung Statistik – Ergebniskontrolle

1.
   - Modus: $x_{Mo} = 4$ Semester
   - Median: $\tilde{x} = 6$ Semester
   - arithmetisches Mittel: $\bar{x} = 7,23$ Semester
   - Spannweite: $V = 17$ Semester
   - $SAQ = 266,31$; $SAQ/n = s^2 = 20,49$
   - $s = \sqrt{s^2} = 4,53$
   - Die Verteilung ist linkssteil (=rechtssschief): die meisten Werte (Modus) liegen links, deshalb ist der Modus der niedrigste der drei Mittelwerte, dann folgt der Median, dann das aritmetische Mittel

2.
   - Republikaner insgesamt: $100/2200 = 4,55\%$; Männer: $70/1000 = 7\%$; Frauen: $30/1200 = 2,5\%$
   - Prozentpunkte: $7 - 2,5 = 4,5$ Prozentpunkte; Um wieviel Prozent höher: $\frac{\text{Differenz}}{\text{Basis}} = \frac{4,5}{2,5} = 1,8 = 180\%$
   - $\chi^2 = 25,46$; $V = 0,11$
   - Ausgeschlossen wird der α-Fehler, d.h. die Wahrscheinlichkeit, die Nullhypothese zu unrecht zu verwerfen ($\alpha \leq 0,001$)

3.
   - Streudiagramm: Bitte selbst malen!
   - $\hat{y} = a + b \times x = -110,31 + 1,01 \times x$
   - Die Konstante entspricht dem Achsenabschnitt, d.h. dem Gewicht, das für eine Person mit einer Körperlänge von 0cm geschätzt würde – keine sehr anschauliche Interpretation
   - Der Koeffizient entspricht der Zunahme des geschätzten Gewichtes pro zusätzlichen Zentimeter Körperlänge (eine Person B, die einen Zentimeter größter ist als eine Person A, sollte nach dem Modell 1,01 Kilo mehr wiegen)
   - Bei einer waagerechten Geraden wäre der Koeffizient gleich null, d.h. das Gewicht wäre von der Körpergröße unabhängig.
   - $R^2 = 0,94$, d.h. 94% der Varianz des Gewichtes können durch das Modell erklärt/vorhergesagt werden
   - $\hat{y} = -110,31 + 1,01 \times 210 = 101,79$
   - $r = \sqrt{R^2} = 0.97$

4.
   - SPD-Anteil an den Zweitstimmen 1990 in Bayern: $1698/6367,2 = 26,67\%$; 1994: $1984/6693 = 29,64\%$

- Differenz: $29,64 - 26,67 = 2,97$ Prozentpunkte; prozentualter Zuwachs: $2,97/26,67 = 11,14\%$
- Für die anderen Bundesländer gehen Sie bitte analog vor

5. -
     - Berechnung des Standardfehlers: $\sqrt{\frac{4\% \times 96\%}{1000}} = 0,62\%$
     - kritischer z-Wert, der symmetrisch 1% einer Standardnormalverteilung abtrennt: $\pm 2,58$
     - Intervall: $4\% - 2,58 \times 0,62\% = 2,40\%$;
       $4\% + 2,58 \times 0,62\% = 5,60\%$;
       $2,40\% \leq \Theta \leq 5,60\%$
   - Wenn sich der Stichprobenumfang vervierfacht, halbiert sich der Standardfehler (die Fallzahl steht im Nenner unter der Wurzel). Damit halbiert sich auch die Breite des Konfidenzintervalls.
   - Das bekommen Sie jetzt selbst hin!

6. Frage an Sie: Haben Sie wirklich versucht, das auszurechnen??? Nennen Sie zwei Gründe, warum Sie das nicht können und sollen!

7. $\tilde{x} = 37; \bar{x} = 42; s^2 = 341,65; \hat{\sigma^2} = 363$

8. $V = 0,15; \chi^2 = 64,76; df = 4; p \leq 0,001$

9. - Einstufige Zufallsauswahl
   - Standardfehler: $\frac{15}{\sqrt{1200}} = 0,43$
   - Intervall: $95 \pm 2,58 \times 0,43$;
     $93,89 \leq \mu \leq 96,11$
   - Stichprobenumfang und Streuung in der Grundgesamtheit

10. - $(70 + 4)/200 = 37\%$
    - γ
    - $N_C = 136084; N_D = 10342; \gamma = \frac{N_C - N_D}{N_C + N_D} = \frac{125742}{146426} = 0,86$

11. - $\bar{x} = 3,88$
    - $s^2 = 1,79; s = 1,34$

12. - lineare Regression
    - $\hat{y} = a + b \times x = 5,43 - 0,20x$
    - $r = -0,48; r^2 = 0,23$
    - $5,23$

13. - $628/850 = 74\%$ der CDU-Wähler wollen die AKWs weiterbetreiben, aber nur $322/800 = 40,2\%$ der SPD-Wähler etc.

- $V = 0,34$
- $\chi^2 = 190,85$ bei $df = 1 \ldots$

14.
- Wahlbeteiligung $= \frac{\text{gültige Stimmen+ungültige Stimmen}}{\textit{Wahlberechtigte}}$
- Wahlbeteiligung für Bayern: $\frac{6367{,}2+52{,}6}{8623{,}6} = 74,44\%$
- Anteil gültiger Stimmen in Bayern: $\frac{6367{,}2}{6367{,}2+52{,}6} = 99,18\%$
- Veränderungen in Prozenten / Prozentpunkte vgl. Aufgabe Nr. 4
- Berechnung für die anderen Länder analog

15.
- Vorgehen wie bei Aufgabe 9; Standardfehler = 0,03; kritischer Wert 1,96; Intervall $3,23 \leq \mu \leq 3,37$
- Intervall wird breiter: kritischer Wert 2,58; Intervall $3,21 \leq \mu \leq 3,39$

16.
- Abhängig vom Skalenniveau der untersuchten Variablen; $V; \lambda; \gamma; \eta; \eta^2; r;$
- positiver, mittelmäßig starker Zusammenhang; beide Variablen haben $0,56^2 = 31\%$ Varianz gemeinsam.
- Um zu ermitteln, wie sich die abhängige Variable auf Veränderungen der unabhängigen reagiert