

PROBEKLAUSUR STATISTIK I

BESPRECHUNG DER KLAUSUR AB: 18.7.2006
HILFSMITTEL: A4 BLATT MIT NOTIZEN, TR

Die Prüfung besteht aus zwei Teilen, die insgesamt 300 Punkte ergeben:

Teil 1: Multiple Choice (jede Aufgabe 5 Punkte)

Teil 2: 5 von 8 Aufgaben mit jeweils 50 Punkten

Bitte ankreuzen, welche Aufgaben bewertet werden sollen (5 von 8)!

| Aufgabe: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | MC | Gesamt |
|-----------|---|---|---|---|---|---|---|---|----|--------|
| Bewerten: | | | | | | | | | | |
| Punkte: | | | | | | | | | | |

Schreiben Sie bitte auf jedes Blatt oben rechts Ihren Namen!

Viel Erfolg!

Teil 1. : Multiple Choice

Zu jeder Frage ist genau eine richtige Antwortmöglichkeit vorgegeben. Tragen Sie Ihre Lösungen in die Kästchen auf der **übernächsten** Seite ein. Die Rückseite der Blätter können Sie für Berechnungen sowie zu Anmerkungen und Erläuterungen Ihrer Lösung verwenden.

- (1) Welche Aussage über einen Box-Plot ist richtig? Aus einem Box-Plot kann man
 - (a) Minimum, unteres Quartil, Median, oberes Quartil, Maximum
 - (b) Minimum, unteres Quantil, Median, oberes Quantil, Maximum
 - (c) Minimum, unteren Hinge, Mittelwert, oberen Hinge, Maximum
 - (d) Minimum, unteren Hinge, Median, oberen Hinge, Maximum
 ablesen.

- (2) Welche Aussage über ML-Schätzer ist richtig?
 - (a) ML-Schätzer sind immer erwartungstreu.
 - (b) ML-Schätzer sind konsistent.
 - (c) ML-Schätzer haben die größte Effizienz bei normalverteilten Zufallsvariablen.
 - (d) ML-Schätzer können nur berechnet werden, wenn die Dichte der ZVs zweimal stetig differenzierbar sind.

- (3) Welche der folgenden Aussagen über den zweiseitigen Test der Hypothese $H_0 : \theta = \theta_0$ gegen $H_1 : \theta \neq \theta_0$ ist wahr?
- (a) Je größer der p -Wert desto stärker sprechen die Daten gegen die Nullhypothese.
 - (b) Die Teststatistik $|T|$ für den zweiseitigen Test wird mit dem $(1 - \alpha)$ - Quantil der Verteilung von T verglichen.
 - (c) Ein p -Wert kleiner als 0.06 führt zu einer Ablehnung der Nullhypothese zum Niveau $\alpha = 0.1$.
 - (d) Ein p -Wert kleiner als 0.05 besagt, daß θ_0 innerhalb eines 95% Konfidenzintervalls für θ liegt.
- (4) Welche der folgenden Aussagen zählt nicht zu den Standardvoraussetzungen des klassischen linearen Modelles?
- (a) Die Beobachtungen sind unkorreliert.
 - (b) Die erklärende Variable X ist normalverteilt.
 - (c) Die Fehler sind identisch verteilt mit Erwartungswert 0.
 - (d) Die Standardabweichung der Fehler hängt nicht von der erklärenden Variable X ab.
- (5) Welche der folgenden Aussagen über einen χ^2 -Test ist richtig?
- (a) Um den χ^2 -Test durchzuführen, benötigt man eine χ^2 -Verteilung mit 4 Freiheitsgraden.
 - (b) Die Nullhypothese und die Alternativhypothese sind für alle χ^2 -Tests gleich.
 - (c) Der χ^2 -Test bewertet nur die Abhängigkeit von zwei Variablen und ist von der Stichprobengröße unabhängig.
 - (d) Die Anzahl der Freiheitsgrade der χ^2 -Verteilung wächst mit der Anzahl der Klassen der Variablen.
- (6) Eine Geschwindigkeitsmessung bei einer Verkehrskontrolle in einer 30 km/h - Zone ergab folgende Werte (in km/h): 45, 33, 28, 23, 41, 35, 58, 34, 31, 45. Welche Aussage über die Kenngrößen ist falsch?
- (a) Der Median ist 34.
 - (b) Die Spannbreite der Daten ist 35.
 - (c) Der Modus ist 45.
 - (d) Im Schnitt wurde mehr als 5 km/h zu schnell gefahren.
- (7) Sei eine a-posteriori Verteilung $Beta(a, b)$ gegeben. Wir fügen nun Information aus einer binomial verteilten Stichprobe $B(N, p)$ mit x Treffern hinzu. Dann ist die a-posteriori Verteilung:
- (a) Normalverteilt mit Parametern a und b .
 - (b) Beta verteilt mit Parametern a und b .
 - (c) Beta verteilt mit Parametern $a' = a + x$ und $b' = b + N - x$.
 - (d) Binomial verteilt mit Parametern $a' = a + x$ und $b' = b + N - x$.

Verteilungstabelle der Standardnormalverteilung

| x | $\Phi(x)$ | x | $\Phi(x)$ | x | $\Phi(x)$ | x | $\Phi(x)$ |
|-------|-----------|-------|-----------|-------|-----------|-------|-----------|
| 0.000 | 0.5000 | 0.750 | 0.7734 | 1.500 | 0.9332 | 2.250 | 0.9878 |
| 0.025 | 0.5100 | 0.775 | 0.7808 | 1.525 | 0.9364 | 2.275 | 0.9885 |
| 0.050 | 0.5199 | 0.800 | 0.7881 | 1.550 | 0.9394 | 2.300 | 0.9893 |
| 0.075 | 0.5299 | 0.825 | 0.7953 | 1.575 | 0.9424 | 2.325 | 0.9900 |
| 0.100 | 0.5398 | 0.850 | 0.8023 | 1.600 | 0.9452 | 2.350 | 0.9906 |
| 0.125 | 0.5497 | 0.875 | 0.8092 | 1.625 | 0.9479 | 2.375 | 0.9912 |
| 0.150 | 0.5596 | 0.900 | 0.8159 | 1.650 | 0.9505 | 2.400 | 0.9918 |
| 0.175 | 0.5695 | 0.925 | 0.8225 | 1.675 | 0.9530 | 2.425 | 0.9923 |
| 0.200 | 0.5793 | 0.950 | 0.8289 | 1.700 | 0.9554 | 2.450 | 0.9929 |
| 0.225 | 0.5890 | 0.975 | 0.8352 | 1.725 | 0.9577 | 2.475 | 0.9933 |
| 0.250 | 0.5987 | 1.000 | 0.8413 | 1.750 | 0.9599 | 2.500 | 0.9938 |
| 0.275 | 0.6083 | 1.025 | 0.8473 | 1.775 | 0.9621 | 2.525 | 0.9942 |
| 0.300 | 0.6179 | 1.050 | 0.8531 | 1.800 | 0.9641 | 2.550 | 0.9946 |
| 0.325 | 0.6274 | 1.075 | 0.8588 | 1.825 | 0.9660 | 2.575 | 0.9950 |
| 0.350 | 0.6368 | 1.100 | 0.8643 | 1.850 | 0.9678 | 2.600 | 0.9953 |
| 0.375 | 0.6462 | 1.125 | 0.8697 | 1.875 | 0.9696 | 2.625 | 0.9957 |
| 0.400 | 0.6554 | 1.150 | 0.8749 | 1.900 | 0.9713 | 2.650 | 0.9960 |
| 0.425 | 0.6646 | 1.175 | 0.8800 | 1.925 | 0.9729 | 2.675 | 0.9963 |
| 0.450 | 0.6736 | 1.200 | 0.8849 | 1.950 | 0.9744 | 2.700 | 0.9965 |
| 0.475 | 0.6826 | 1.225 | 0.8897 | 1.975 | 0.9759 | 2.725 | 0.9968 |
| 0.500 | 0.6915 | 1.250 | 0.8944 | 2.000 | 0.9772 | 2.750 | 0.9970 |
| 0.525 | 0.7002 | 1.275 | 0.8988 | 2.025 | 0.9786 | 2.775 | 0.9972 |
| 0.550 | 0.7088 | 1.300 | 0.9032 | 2.050 | 0.9798 | 2.800 | 0.9974 |
| 0.575 | 0.7174 | 1.325 | 0.9074 | 2.075 | 0.9810 | 2.825 | 0.9976 |
| 0.600 | 0.7257 | 1.350 | 0.9115 | 2.100 | 0.9821 | 2.850 | 0.9978 |
| 0.625 | 0.7340 | 1.375 | 0.9154 | 2.125 | 0.9832 | 2.875 | 0.9980 |
| 0.650 | 0.7422 | 1.400 | 0.9192 | 2.150 | 0.9842 | 2.900 | 0.9981 |
| 0.675 | 0.7502 | 1.425 | 0.9229 | 2.175 | 0.9852 | 2.925 | 0.9983 |
| 0.700 | 0.7580 | 1.450 | 0.9265 | 2.200 | 0.9861 | 2.950 | 0.9984 |
| 0.725 | 0.7658 | 1.475 | 0.9299 | 2.225 | 0.9870 | 2.975 | 0.9985 |

Verteilungstabelle der Poissonverteilung

| $\lambda=15$ | | $\lambda=7$ | | $\lambda=6$ | | $\lambda=3$ | |
|--------------|---------------|-------------|---------------|-------------|---------------|-------------|---------------|
| x | $P(X \leq x)$ | x | $P(X \leq x)$ | x | $P(X \leq x)$ | x | $P(X \leq x)$ |
| 1 | 0 | 0.5 | 0.0009 | 0.5 | 0.0025 | 0.2 | 0.0498 |
| 2 | 0 | 1 | 0.0073 | 1 | 0.0174 | 0.4 | 0.0498 |
| 3 | 0.0002 | 1.5 | 0.0073 | 1.5 | 0.0174 | 0.6 | 0.0498 |
| 4 | 0.0009 | 2 | 0.0296 | 2 | 0.062 | 0.8 | 0.0498 |
| 5 | 0.0028 | 2.5 | 0.0296 | 2.5 | 0.062 | 1 | 0.1991 |
| 6 | 0.0076 | 3 | 0.0818 | 3 | 0.1512 | 1.2 | 0.1991 |
| 7 | 0.018 | 3.5 | 0.0818 | 3.5 | 0.1512 | 1.4 | 0.1991 |
| 8 | 0.0374 | 4 | 0.173 | 4 | 0.2851 | 1.6 | 0.1991 |
| 9 | 0.0699 | 4.5 | 0.173 | 4.5 | 0.2851 | 1.8 | 0.1991 |
| 10 | 0.1185 | 5 | 0.3007 | 5 | 0.4457 | 2 | 0.4232 |
| 11 | 0.1848 | 5.5 | 0.3007 | 5.5 | 0.4457 | 2.2 | 0.4232 |
| 12 | 0.2676 | 6 | 0.4497 | 6 | 0.6063 | 2.4 | 0.4232 |
| 13 | 0.3632 | 6.5 | 0.4497 | 6.5 | 0.6063 | 2.6 | 0.4232 |
| 14 | 0.4657 | 7 | 0.5987 | 7 | 0.744 | 2.8 | 0.4232 |
| 15 | 0.5681 | 7.5 | 0.5987 | 7.5 | 0.744 | 3 | 0.6472 |
| 16 | 0.6641 | 8 | 0.7291 | 8 | 0.8472 | 3.2 | 0.6472 |
| 17 | 0.7489 | 8.5 | 0.7291 | 8.5 | 0.8472 | 3.4 | 0.6472 |
| 18 | 0.8195 | 9 | 0.8305 | 9 | 0.9161 | 3.6 | 0.6472 |
| 19 | 0.8752 | 9.5 | 0.8305 | 9.5 | 0.9161 | 3.8 | 0.6472 |
| 20 | 0.917 | 10 | 0.9015 | 10 | 0.9574 | 4 | 0.8153 |
| 21 | 0.9469 | 10.5 | 0.9015 | 10.5 | 0.9574 | 4.2 | 0.8153 |
| 22 | 0.9673 | 11 | 0.9467 | 11 | 0.9799 | 4.4 | 0.8153 |
| 23 | 0.9805 | 11.5 | 0.9467 | 11.5 | 0.9799 | 4.6 | 0.8153 |
| 24 | 0.9888 | 12 | 0.973 | 12 | 0.9912 | 4.8 | 0.8153 |
| 25 | 0.9938 | 12.5 | 0.973 | 12.5 | 0.9912 | 5 | 0.9161 |
| 26 | 0.9967 | 13 | 0.9872 | 13 | 0.9964 | 5.2 | 0.9161 |
| 27 | 0.9983 | 13.5 | 0.9872 | 13.5 | 0.9964 | 5.4 | 0.9161 |
| 28 | 0.9991 | 14 | 0.9943 | 14 | 0.9986 | 5.6 | 0.9161 |
| 29 | 0.9996 | 14.5 | 0.9943 | 14.5 | 0.9986 | 5.8 | 0.9161 |
| 30 | 0.9998 | 15 | 0.9976 | 15 | 0.9995 | 6 | 0.9665 |
| 31 | 0.9999 | 15.5 | 0.9976 | 15.5 | 0.9995 | 6.2 | 0.9665 |
| 32 | 1 | 16 | 0.999 | 16 | 0.9998 | 6.4 | 0.9665 |
| 33 | 1 | 16.5 | 0.999 | 16.5 | 0.9998 | 6.6 | 0.9665 |
| 34 | 1 | 17 | 0.9996 | 17 | 0.9999 | 6.8 | 0.9665 |
| 35 | 1 | 17.5 | 0.9996 | 17.5 | 0.9999 | 7 | 0.9881 |

Quantile $\chi_{df;1-\alpha}^2$ der χ^2 -Verteilung

| df \ α | 0.100 | 0.050 | 0.025 | 0.020 | 0.010 | 0.005 | 0.0025 | 0.001 | 0.0005 |
|---------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 | 2.705 | 3.841 | 5.023 | 5.411 | 6.634 | 7.879 | 9.140 | 10.827 | 12.115 |
| 2 | 4.605 | 5.991 | 7.377 | 7.824 | 9.210 | 10.596 | 11.982 | 13.815 | 15.201 |
| 3 | 6.251 | 7.814 | 9.348 | 9.837 | 11.344 | 12.838 | 14.320 | 16.266 | 17.730 |
| 4 | 7.779 | 9.487 | 11.143 | 11.667 | 13.276 | 14.860 | 16.423 | 18.466 | 19.997 |
| 5 | 9.236 | 11.070 | 12.832 | 13.388 | 15.086 | 16.749 | 18.385 | 20.515 | 22.105 |
| 6 | 10.644 | 12.591 | 14.449 | 15.033 | 16.811 | 18.547 | 20.249 | 22.457 | 24.102 |
| 7 | 12.017 | 14.067 | 16.012 | 16.622 | 18.475 | 20.277 | 22.040 | 24.321 | 26.017 |
| 8 | 13.361 | 15.507 | 17.534 | 18.168 | 20.090 | 21.954 | 23.774 | 26.124 | 27.868 |
| 9 | 14.683 | 16.918 | 19.022 | 19.679 | 21.665 | 23.589 | 25.462 | 27.877 | 29.665 |
| 10 | 15.987 | 18.307 | 20.483 | 21.160 | 23.209 | 25.188 | 27.112 | 29.588 | 31.419 |
| 11 | 17.275 | 19.675 | 21.920 | 22.617 | 24.724 | 26.756 | 28.729 | 31.264 | 33.136 |
| 12 | 18.549 | 21.026 | 23.336 | 24.053 | 26.216 | 28.299 | 30.318 | 32.909 | 34.821 |
| 13 | 19.811 | 22.362 | 24.735 | 25.471 | 27.688 | 29.819 | 31.883 | 34.528 | 36.477 |
| 14 | 21.064 | 23.684 | 26.118 | 26.872 | 29.141 | 31.319 | 33.426 | 36.123 | 38.109 |
| 15 | 22.307 | 24.995 | 27.488 | 28.259 | 30.577 | 32.801 | 34.949 | 37.697 | 39.718 |

Quantile $t_{n;1-\alpha}$ der t -Verteilung

| $n \setminus \alpha$ | 0.100 | 0.050 | 0.025 | 0.020 | 0.010 | 0.005 | 0.0025 | 0.001 | 0.0005 |
|----------------------|-------|-------|--------|--------|--------|--------|---------|---------|---------|
| 1 | 3.077 | 6.313 | 12.706 | 15.894 | 31.820 | 63.656 | 127.321 | 318.308 | 636.619 |
| 2 | 1.885 | 2.919 | 4.302 | 4.848 | 6.964 | 9.924 | 14.089 | 22.327 | 31.599 |
| 3 | 1.637 | 2.353 | 3.182 | 3.481 | 4.540 | 5.840 | 7.453 | 10.214 | 12.923 |
| 4 | 1.533 | 2.131 | 2.776 | 2.998 | 3.746 | 4.604 | 5.597 | 7.173 | 8.610 |
| 5 | 1.475 | 2.015 | 2.570 | 2.756 | 3.364 | 4.032 | 4.773 | 5.893 | 6.868 |
| 6 | 1.439 | 1.943 | 2.446 | 2.612 | 3.142 | 3.707 | 4.316 | 5.207 | 5.958 |
| 7 | 1.414 | 1.894 | 2.364 | 2.516 | 2.997 | 3.499 | 4.029 | 4.785 | 5.407 |
| 8 | 1.396 | 1.859 | 2.306 | 2.448 | 2.896 | 3.355 | 3.832 | 4.500 | 5.041 |
| 9 | 1.383 | 1.833 | 2.262 | 2.398 | 2.821 | 3.249 | 3.689 | 4.296 | 4.780 |
| 10 | 1.372 | 1.812 | 2.228 | 2.359 | 2.763 | 3.169 | 3.581 | 4.143 | 4.586 |
| 11 | 1.363 | 1.795 | 2.200 | 2.328 | 2.718 | 3.105 | 3.496 | 4.024 | 4.436 |
| 12 | 1.356 | 1.782 | 2.178 | 2.302 | 2.680 | 3.054 | 3.428 | 3.929 | 4.317 |
| 13 | 1.350 | 1.770 | 2.160 | 2.281 | 2.650 | 3.012 | 3.372 | 3.851 | 4.220 |
| 14 | 1.345 | 1.761 | 2.144 | 2.263 | 2.624 | 2.976 | 3.325 | 3.787 | 4.140 |
| 15 | 1.340 | 1.753 | 2.131 | 2.248 | 2.602 | 2.946 | 3.286 | 3.732 | 4.072 |
| 16 | 1.336 | 1.745 | 2.119 | 2.235 | 2.583 | 2.920 | 3.251 | 3.686 | 4.014 |
| 17 | 1.333 | 1.739 | 2.109 | 2.223 | 2.566 | 2.898 | 3.222 | 3.645 | 3.965 |
| 18 | 1.330 | 1.734 | 2.100 | 2.213 | 2.552 | 2.878 | 3.196 | 3.610 | 3.921 |
| 19 | 1.327 | 1.729 | 2.093 | 2.204 | 2.539 | 2.860 | 3.173 | 3.579 | 3.883 |
| 20 | 1.325 | 1.724 | 2.085 | 2.196 | 2.527 | 2.845 | 3.153 | 3.551 | 3.849 |
| 25 | 1.316 | 1.708 | 2.059 | 2.166 | 2.485 | 2.787 | 3.078 | 3.450 | 3.725 |
| 50 | 1.298 | 1.675 | 2.008 | 2.108 | 2.403 | 2.677 | 2.936 | 3.261 | 3.496 |
| 100 | 1.290 | 1.660 | 1.983 | 2.080 | 2.364 | 2.625 | 2.870 | 3.173 | 3.390 |

Verteilungsfunktion der Binomialverteilung $B_{n;p}(k)$ für $p = 0.5$.

| $n \setminus k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 0.500 | 1.000 | | | | | | | | | |
| 2 | 0.250 | 0.750 | 1.000 | | | | | | | | |
| 3 | 0.125 | 0.500 | 0.875 | 1.000 | | | | | | | |
| 4 | 0.062 | 0.313 | 0.687 | 0.938 | 1.000 | | | | | | |
| 5 | 0.031 | 0.188 | 0.500 | 0.812 | 0.969 | 1.000 | | | | | |
| 6 | 0.016 | 0.109 | 0.344 | 0.656 | 0.891 | 0.984 | 1.000 | | | | |
| 7 | 0.008 | 0.063 | 0.227 | 0.500 | 0.773 | 0.938 | 0.992 | 1.000 | | | |
| 8 | 0.004 | 0.035 | 0.145 | 0.363 | 0.637 | 0.855 | 0.965 | 0.996 | 1.000 | | |
| 9 | 0.002 | 0.020 | 0.090 | 0.254 | 0.500 | 0.746 | 0.910 | 0.980 | 0.998 | 1.000 | |
| 10 | 0.001 | 0.011 | 0.055 | 0.172 | 0.377 | 0.623 | 0.828 | 0.945 | 0.989 | 0.999 | 1.000 |

Kritische Werte $W_{n;\alpha}$ des Vorzeichenrangtests von Wilcoxon

| $n \setminus \alpha$ | 0.01 | 0.025 | 0.05 | 0.1 | 0.9 | 0.95 | 0.975 | 0.99 |
|----------------------|------|-------|------|-----|-----|------|-------|------|
| 4 | 0 | 0 | 0 | 1 | 9 | 10 | 10 | 10 |
| 5 | 0 | 0 | 1 | 3 | 12 | 14 | 15 | 15 |
| 6 | 0 | 1 | 3 | 4 | 17 | 18 | 20 | 21 |
| 7 | 1 | 3 | 4 | 6 | 22 | 24 | 25 | 27 |
| 8 | 2 | 4 | 6 | 9 | 27 | 30 | 32 | 34 |
| 9 | 4 | 6 | 9 | 11 | 34 | 36 | 39 | 41 |
| 10 | 6 | 9 | 11 | 15 | 40 | 44 | 46 | 49 |
| 11 | 8 | 11 | 14 | 18 | 48 | 52 | 55 | 58 |
| 12 | 10 | 14 | 18 | 22 | 56 | 60 | 64 | 68 |
| 13 | 13 | 18 | 22 | 27 | 64 | 69 | 73 | 78 |
| 14 | 16 | 22 | 26 | 32 | 73 | 79 | 83 | 89 |
| 15 | 20 | 26 | 31 | 37 | 83 | 89 | 94 | 100 |
| 16 | 24 | 30 | 36 | 43 | 93 | 100 | 106 | 112 |
| 17 | 28 | 35 | 42 | 49 | 104 | 111 | 118 | 125 |
| 18 | 33 | 41 | 48 | 56 | 115 | 123 | 130 | 138 |
| 19 | 38 | 47 | 54 | 63 | 127 | 136 | 143 | 152 |
| 20 | 44 | 53 | 61 | 70 | 140 | 149 | 157 | 166 |

Teil 2

Bearbeiten Sie 5 der 8 Aufgaben!

1. BUNDESLIGA

Es wird oft vorgeschlagen, die Anzahl der Tore in einem Fußballspiel durch eine Poisson-verteilte Zufallsvariable zu beschreiben. Folgende Resultate wurden am 5. Bundesligaspieltag in der Saison 2003/04 der Fußball Bundesliga erzielt:

| | | |
|---------------------|-----------------------|-----|
| TSV 1860 München | : 1.FC Köln | 2:1 |
| Borussia Dortmund | : Werder Bremen | 2:1 |
| Hertha BSC Berlin | : Hannover 96 | 2:3 |
| VfL Wolfsburg | : FC Bayern München | 3:2 |
| Hansa Rostock | : VfL Bochum | 0:2 |
| FC Schalke 04 | : VfB Stuttgart | 0:0 |
| Bayer 04 Leverkusen | : Hamburger SV | 1:0 |
| 1.FC Kaiserslautern | : SC Freiburg | 2:2 |
| Borussia M'gladbach | : Eintracht Frankfurt | 0:2 |

Rudi Völler und Günter Netzer diskutieren über die Ergebnisse. Im Verlauf der Diskussion behauptet Herr Netzer, der zugrundeliegende Poissonparameter sei gleich 3. Herr Völler läuft rot an, und widerspricht. Sie müssen als Mitglied einer unabhängigen Expertenkommission die Streitfrage entscheiden.

- Leiten Sie einen Test für obiges Entscheidungsproblem her.
- Skizzieren Sie die Gütefunktion Ihres Testes.
- Führen Sie den Test für obige Daten durch und teilen Sie das Ergebnis in verständlicher Weise den Herren Völler und Netzer mit.
- Versuchen Sie den Herren Völler und Netzer kurz und knapp die Begriffe Signifikanzniveau, Gütefunktion und Fehler 1. und 2. Art zu erklären. ■

2. UNI ZULASSUNG

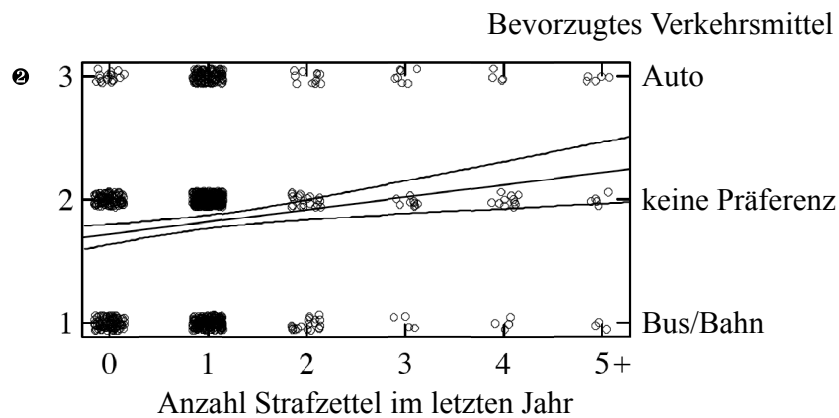
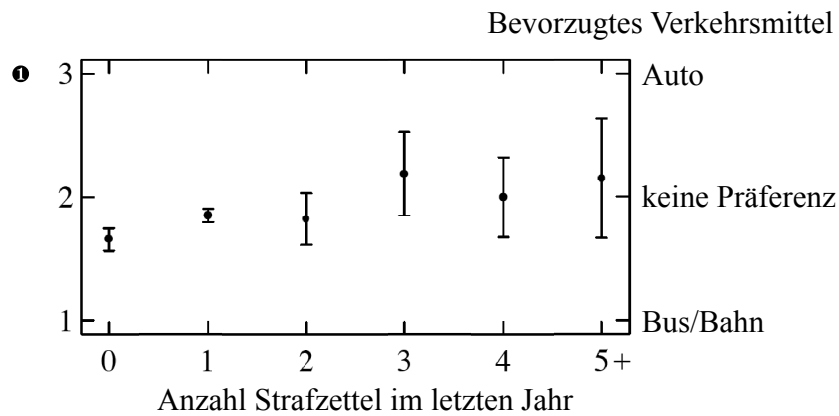
An der Universität von Berkeley haben sich in einem Jahrgang insgesamt 4526 Studenten beworben, davon wurden 1755 aufgenommen. Nach Fachrichtung und Geschlecht aufgeteilt ergaben sich folgende Zulassungszahlen:

| Fachrichtung | Geschlecht | zugelassen? | Anzahl |
|----------------|------------|-------------|--------|
| Ingenieurwesen | männl. | ja | 865 |
| Ingenieurwesen | männl. | nein | 520 |
| Ingenieurwesen | weibl. | ja | 106 |
| Ingenieurwesen | weibl. | nein | 27 |
| andere Fächer | männl. | ja | 333 |
| andere Fächer | männl. | nein | 973 |

- Sie bekamen die Zulassungstabelle per Fax und dabei brach die Verbindung vorzeitig ab. Deswegen wurden leider die letzten Zeilen der Tabelle abgeschnitten. Vervollständigen Sie die Tabelle.
- Der Universität wird vorgeworfen, dass die Geschlechter bei der Zulassung unterschiedlich behandelt werden. Überprüfen Sie anhand der vorliegenden Daten mit einem geeigneten Test auf dem Signifikanzniveau von 5%, ob dies der Fall ist. Lassen Sie dabei die Fachrichtung ausser acht. Falls keine Gleichbehandlung vorliegt, welches Geschlecht wird benachteiligt?
- Die zentrale Universitätsverwaltung ist mit dem Ergebnis unzufrieden und ordnet an, dass jede Fachrichtung einzeln überprüft werden soll. Führen Sie den Test aus (b) für jede Fachrichtung einzeln durch. Was ist das Ergebnis? Welches Geschlecht hat jeweils die höheren Zulassungsquoten?
- Vergleichen Sie die Ergebnisse aus (b) und (c). Die Universität fragt Sie, ob Frauen benachteiligt werden. Beantworten Sie die Frage und erläutern Sie kurz Ihre Antwort.

3. GRAFIK

10.000 Münchener Bürger wurden nach ihrem bevorzugtem Verkehrsmittel und der Anzahl der Strafzettel, die im letzten Jahr gegen sie verhängt wurden befragt. Folgende zwei Grafiken wurden zu den Daten dieser Umfrage veröffentlicht:

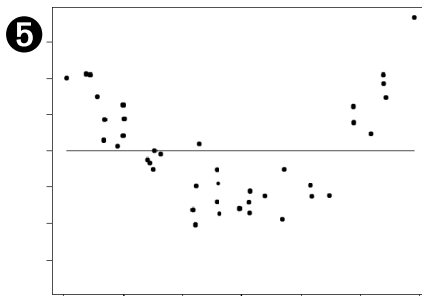
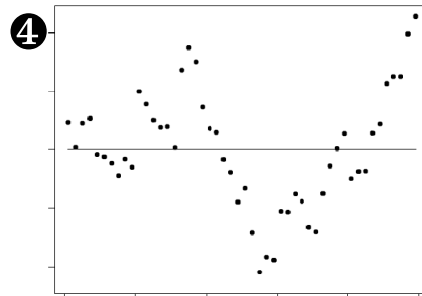
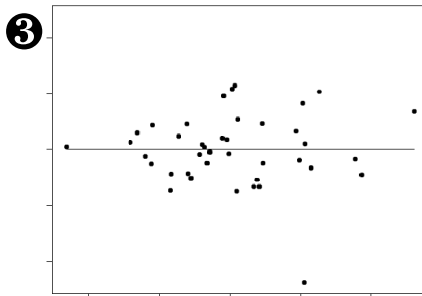
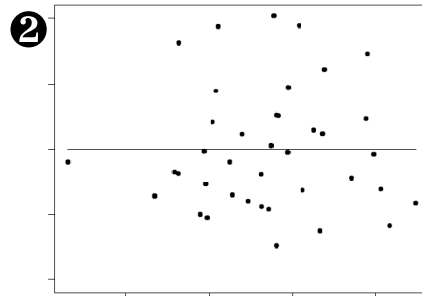
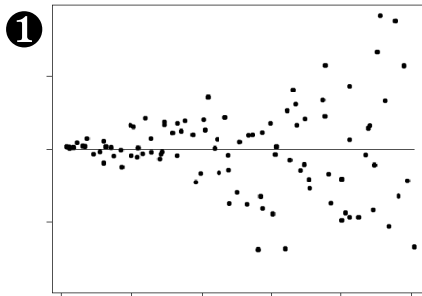


- (a) Beschreiben Sie, wie in den beiden Grafiken die Daten dargestellt werden.
- (b) Kritisieren Sie die beiden Darstellungen.
- (c) Machen Sie zwei Alternativvorschläge, und diskutieren Sie deren Vor- und Nachteile.

4. RESIDUAL ANALYSE

Im einfachen linearen Modell $y_i = \alpha + \beta x_i + \epsilon_i$ sollte man die Qualität der Anpassung über einfache Streudiagramme der Einflussgröße gegen die Residuen überprüfen.

- Welche Annahmen stellt man an die Fehlerterme ϵ_i (bzw. $\hat{\epsilon}_i$)?
- Welche Modellannahmen sind weiterhin zu erfüllen?
- Nachstehend sind fünf Residuen-Plots abgebildet. Beschreiben Sie zu jeder Graphik ob, und wenn ja, welche Voraussetzungen verletzt worden sind.



5. DOSENPFAND

Um die Wirksamkeit des Dosenpfands nach Minister Tritthn zu überprüfen, wurde die durchschnittliche Anzahl der zurückgegebenen Dosen pro Monat pro Geschäft in 10 verschiedenen Vertriebsregionen der Alti Kette aufgenommen.

| | | | | | | | | | | |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Vorher | 119.0 | 114.9 | 131.8 | 118.7 | 106.0 | 118.2 | 126.6 | 124.1 | 126.3 | 117.6 |
| Nachher | 122.3 | 117.3 | 141.0 | 120.1 | 104.6 | 119.1 | 128 | 132.8 | 131.1 | 118.2 |

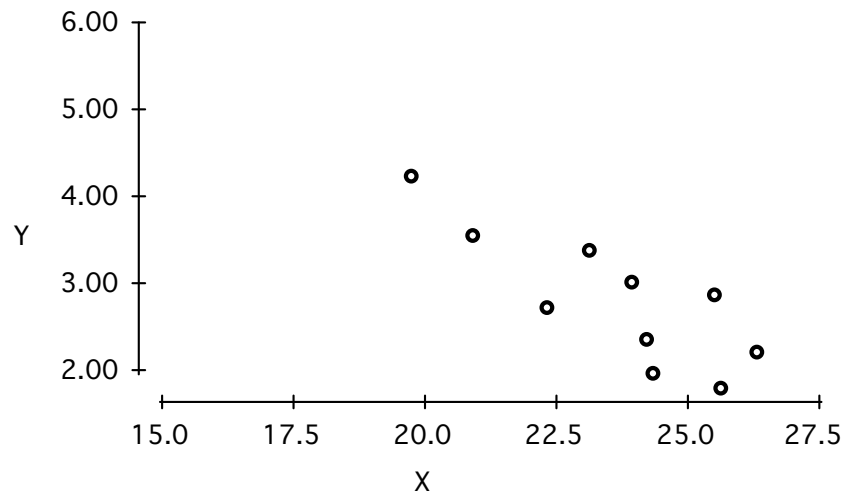
- (a) Erlauben diese Daten die Behauptung, dass nach Einführung des Pfands signifikant mehr Dosen zurückgegeben wurden (ein Niveau von $\alpha = 0.1$ sei vorausgesetzt), bei Verwendung eines
- t-Tests
 - geeigneten nicht parametrischen Tests.
- (b) Angenommen die 20 Werte kommen aus 20 verschiedenen Regionen.
- Welcher nicht parametrische Test ist dann anzuwenden? Berechnen Sie die entsprechende Teststatistik.
 - Zu welchem Niveau kann der Dosenpfand in diesem Fall als Erfolg gewertet werden? Benutzen Sie hier die Normalapproximation.

6. CD-SPINDEL

Assistent U brennt seine CDs immer nur mit 48-facher Geschwindigkeit. In der Zeitschrift c't findet er folgende Ausschussraten (Y) in % und Straßenpreise (X) in EUR für 100er Spindeln unter oben genannten Bedingungen,

| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| (X) | 22.30 | 20.90 | 25.60 | 15.60 | 19.70 | 23.10 | 24.30 | 24.20 | 23.90 | 25.50 | 26.30 |
| (Y) | 2.73 | 3.56 | 1.81 | • | 4.25 | 3.391 | 1.98 | 2.36 | 3.03 | 2.89 | 2.22 |

wobei die Rate für Nummer 4 durch einen Kaffeefleck leider unleserlich ist. Die restlichen Daten trägt Assistent U in folgendes Streudiagramm ein:



(a) Berechnen Sie unter Verwendung von

$$\sum_{i \neq 4} x_i = 235.8, \quad \sum_{i \neq 4} x_i^2 = 5600.44, \quad \sum_{i \neq 4} y_i = 28.22, \quad \sum_{i \neq 4} x_i y_i = 653.377$$

Steigung und Achsenabschnitt des KQ Schätzers der einfachen linearen Regression von X und Y und tragen Sie diesen auch in obiges Streudiagramm ein!

- (b) Geben Sie eine Prognose \hat{Y}_4 für die fehlende Ausschussrate Y_4 an. Beurteilen Sie die Aussagekraft des Schätzers in der gegebenen Situation.
- (c) Mit welchem Ausschuss muss Assistent U rechnen, wenn er eine Spindel der Marke Lowscreen bei Wobis ($i = 4$) kauft. Er möchte eine Sicherheit von 95% für das Intervall. (Hinweis: Die geschätzte Varianz beträgt $\hat{\sigma}^2 = 0.208$)
- (d) Angenommen es gilt $\hat{Y}_4 = Y_4$. Wie ändert sich die Schätzung der Regression für alle 11 Werte. Keine Rechnung erforderlich, nur eine formale Begründung.

7. ERWARTUNGSTREUER SCHÄTZER

Sei $X_i, i = 1, 2, \dots, n$, eine zufällige Stichprobe aus $X \sim N(\theta; 1)$.

- (a) Berechnen Sie die log-Likelihood.
- (b) Berechnen Sie explizit den ML Schätzer für θ .
- (c) Wie sieht der Schätzer aus, falls $X \sim Exp(\theta)$ gilt?

Vergleichen Sie nun, den Schätzer aus (b) mit den folgenden 2 alternativen Schätzern:

- i) $T_{2n} = \frac{2}{n(n+1)} \sum(iX_i)$
- ii) $T_{3n} = \sum(-1)^i X_i$

- (d) Welche der drei Schätzer sind biased? Wie groß ist der Bias?
- (e) Welche Eigenschaft eines Schätzers hängt von der Varianz ab? Welcher der 3 Schätzern ist diesbezüglich am besten? Rechnen Sie die Varianzen der Schätzer aus.

8. LKW-MAUT

Der Fehler der gemessenen Fahrstrecke eines LKWs (pro 100km) sei normalverteilt mit Erwartungswert μ und Varianz σ^2 . In einer Stichprobe von $n = 19$ LKW ergaben sich ein Mittelwert von $\bar{x} = 2$. Der Verband der deutschen Spediteure möchte der Betreiberfirma mit einer Sicherheit von 99.5% nachweisen, dass sie generell zu viele Kilometer abrechnet.

- (a) Wie muss die Hypothese für einen entsprechenden statistischen Test formuliert werden?
- (b) Wie ist das Testergebnis, wenn wir eine bekannte Varianz σ^2 von 9 voraussetzen? Geben Sie den entsprechenden p -Wert an.
- (c) Angenommen wir müssen die Varianz auch empirisch bestimmen, und es ergibt sich $s^2 = 9$. Was ändert sich am Test und dem Ergebnis?
- (d) In einem zweiten Versuch ergibt sich ein \bar{x} von 1 bei der Untersuchung von 64 Fahrten. Wie ändern sich die Ergebnisse aus (b) und (c)?