Lehrstuhl für Ökonometrie und Statistik

Statistik II
Wintersemester 2013/14

Prof. Dr. Wolf-Dieter Heller
Frieder Conrad
Hartwig Senska

Aufgabensammlung zur Vorlesung Statistik II

Inhaltsverzeichnis

1 Informationen zu Vorlesung und Übungsbetrieb

Vorlesung: Die Vorlesung findet

- freitags, 8:00 - 9:30 Uhr im Gaede, Geb. 30.22, sowie
- freitags, 14:00 - 15:30 Uhr im Gaede, Geb. 30.22

statt.

Tutorien: Begleitend zur Vorlesung werden Tutorien angeboten. Ziel der Tutorien ist die Besprechung einer Auswahl von Aufgaben aus dieser Aufgabensammlung. Die Lösungen zu den Aufgaben aus dieser Aufgabensammlung stehen auf der Homepage des Lehrstuhls zum Download bereit. Zu den Tutorien ist keine Anmeldung nötig.

PC-Praktikum: Ergänzend zur Vorlesung wird zu den folgenden Terminen ein PC-gestütztes Praktikum (3 Lektionen à 2 Stunden) im Raum CIP-I des CIP-Pool der Fakultät für Wirtschaftswissenschaften (Sockelgeschoss von Geb. 11.40) angeboten.

- 1. Durchgang: 2., 4. und 9. Dezember 2013, Anmeldung ab dem 25. November 2013
- 2. Durchgang: 8., 13. und 15. Januar 2014, Anmeldung ab dem 16. Dezember 2013
- 3. Durchgang: 22., 27. und 29. Januar 2014, Anmeldung ab dem 15. Januar 2014

Die Zeiten an den einzelnen Tagen sind jeweils 08:00 - 10:00 Uhr, 10:00 - 12:00 Uhr oder 12:00 - 14:00 Uhr. Eine Anmeldung zu den einzelnen Lektionen ist durch Eintrag in die Listen am Schwarzen Brett des Lehrstuhls vor Zi. 214, Geb. 20.12 möglich. Für die erfolgreiche Teilnahme an mindestens fünf der insgesamt sechs Lektionen des PC-Praktikums in Statistik I/II erhält man einen Schein.

Für eine Teilnahme an dem PC-Praktikum werden Grundkenntnisse im Umgang mit Microsoft Excel vorausgesetzt.

Klausur: Die Klausur zur Vorlesung Statistik II findet am Samstag, den 15. Februar 2014 von 8:00 - 10:00 Uhr statt.

Die Anmeldung zur Klausur erfolgt über die Selbstbedienungsfunktion des Studierendenportals. Angehörige anderer Fakultäten, die einen Statistik II Schein erwerben möchten, melden sich mit einem Zettel an, auf dem Name, Matrikelnummer und Studienfach vermerkt sind. Die Anmeldung erfolgt durch die Abgabe dieses Zettels im Sekretariat des Lehrstuhls (Zi. 209 Geb. 20.12). Die Hörsaalverteilungen sind zu gegebener Zeit auf der Prüfungs-Homepage des Lehrstuhls abrufbar.

Die Nachklausur Statistik II findet in der ersten Maihälfte 2014 statt. Der genaue Termin wird rechtzeitig bekannt gegeben.

Beide Klausuren werden ohne Hilfsmittel (mit Ausnahme der vom Lehrstuhl zugelassenen Taschenrechner) geschrieben. Den Klausurexemplaren werden als Hilfestellung eine kleine Formelsammlung sowie die benötigten Tabellen beigelegt. Die Hilfestellungen werden rechtzeitig vor dem Klausurtermin bekannt gegeben. In der Woche vor der Hauptklausur findet eine Fragestunde statt.

Die Ergebnisse aller Klausuren werden lediglich auf der Homepage des Lehrstuhl veröffentlicht und können dort abgefragt werden. Benötigt wird dazu die während der Klausur ausgehändigte PIN!

Ansprechpartner: Inhaltliche Fragen können in den Tutorien und im Ilias-Forum der Veranstaltung gestellt werden. Ihre direkten Ansprechpartner sind Frieder Conrad für Fragen zum Vorlesungs- und Übungsbetrieb und Hartwig Senska für Fragen zur Klausurorganisation. Beide erreichen Sie per Email unter statistik@econ.kit.edu. Für das Rechnerpraktikum ist Florian Jacob zuständig. Die aktuellen Sprechstunden aller Lehrstuhlmitarbeiter entnehmen Sie bitte der Homepage des Lehrstuhls.

Sonstige Informationen: Weitere Informationen werden in der Vorlesung, in den Tutorien, im Ilias-Kurs der Veranstaltung sowie zusätzlich auf den Lehrstuhl-WWW-Seiten (http://statistik.econ.kit.edu) bekannt gegeben.

2 Bücher zur schließenden Statistik

1. Entscheidungstheorie

   - Bamberg, G.; Coenenberg, A.G.: Betriebswirtschaftliche Entscheidungslehre, Vahlen, 14. Aufl., 2008

2. Induktive Statistik

   - Bamberg, G., Baur, F. und Krapp, M.: Statistik, 15. überarb. Auflage. Oldenbourg, München 2011
   - Bol, G.: Induktive Statistik, 3. überarb. Auflage, Oldenbourg, München 2003
   - Bol, G.: Wahrscheinlichkeitstheorie, 6. überarb. Auflage, Oldenbourg, München 2007
   - Mosler, K. und Schmid, F.: Wahrscheinlichkeitsrechnung und schließende Statistik, 4. verb. Aufl., Springer, Berlin 2010
   - Schwarze, J.: Grundlagen der Statistik 2: Wahrscheinlichkeitsrechnung und induktive Statistik, 9. vollst. überarb. Aufl., NWB, Herne 2009

3. Ergänzende Literatur

   - Bosch, K.: Statistik-Taschenbuch, 3. Aufl., Oldenbourg, München etc., 1998
   - Hartung, J., Elpert, B., Klösener, K.-H.: Statistik: Lehr- und Handbuch der angewandten Statistik, 14. Aufl., Oldenbourg, München etc., 2009
   - Rinne, H.: Taschenbuch der Statistik, 4. überarb. u. erw. Auflage, Harri Deutsch, Frankfurt a. M. 2008

3 Verteilungsfunktion der Standardnormalverteilung

Standardnormalverteilung: $\Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5078	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8600	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9065	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9807	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9867	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9958	0.9959	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9993	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
4.0	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
4.4	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

4 Quantile der χ^2-Verteilung

$\chi^2(n)_\alpha$ ist das α-Quantil der $\chi^2(n)$-Verteilung: $\quad F_{\chi^2(n)}(\chi^2(n)_\alpha) = \alpha$

α / n	0.5%	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%	99.5%
1	0.00	0.00	0.00	0.00	0.02	0.45	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	1.39	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	2.37	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	5.35	10.64	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.68	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	14.34	22.31	24.99	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.56	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.50	37.46	40.47	43.19	46.46	59.34	74.39	79.08	83.30	88.40	91.98
70	43.25	45.42	48.75	51.74	55.33	69.34	85.52	90.53	95.03	100.44	104.24
80	51.14	53.52	57.15	60.39	64.28	79.34	96.57	101.88	106.63	112.34	116.35
90	59.17	61.74	65.64	69.13	73.29	89.33	107.60	113.14	118.14	124.13	128.32
100	67.30	70.05	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.82	140.19

5 Quantile der t-Verteilung

$t(n)_\alpha$ ist das α-Quantil der $t(n)$-Verteilung: $\quad F_{t(n)}(t(n)_\alpha) = \alpha$

α / n	90%	95%	97.5%	99%	99.5%
1	3.0776	6.3137	12.7062	31.8205	63.6568
2	1.8856	2.9200	4.3027	6.9646	9.9248
3	1.6377	2.3534	3.1824	4.5407	5.8409
4	1.5332	2.1318	2.7764	3.7469	4.6041
5	1.4759	2.0150	2.5706	3.3649	4.0321
6	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.4149	1.8946	2.3646	2.9980	3.4995
8	1.3968	1.8595	2.3060	2.8965	3.3554
9	1.3830	1.8331	2.2622	2.8214	3.2498
10	1.3721	1.8125	2.2281	2.7638	3.1693
11	1.3634	1.7959	2.2010	2.7181	3.1058
12	1.3562	1.7823	2.1788	2.6810	3.0545
13	1.3502	1.7709	2.1604	2.6503	3.0123
14	1.3450	1.7613	2.1448	2.6245	2.9768
15	1.3406	1.7531	2.1314	2.6025	2.9467
16	1.3368	1.7459	2.1199	2.5835	2.9208
17	1.3334	1.7396	2.1098	2.5669	2.8982
18	1.3304	1.7341	2.1009	2.5524	2.8784
19	1.3277	1.7291	2.0930	2.5395	2.8609
20	1.3253	1.7247	2.0860	2.5280	2.8453
21	1.3232	1.7207	2.0796	2.5176	2.8314
22	1.3212	1.7171	2.0739	2.5083	2.8188
23	1.3195	1.7139	2.0687	2.4999	2.8073
24	1.3178	1.7109	2.0639	2.4922	2.7969
25	1.3163	1.7081	2.0595	2.4851	2.7874
26	1.3150	1.7056	2.0555	2.4786	2.7787
27	1.3137	1.7033	2.0518	2.4727	2.7707
28	1.3125	1.7011	2.0484	2.4671	2.7633
29	1.3114	1.6991	2.0452	2.4620	2.7564
30	1.3104	1.6973	2.0422	2.4573	2.7500
34	1.3070	1.6909	2.0322	2.4412	2.7284
40	1.3031	1.6839	2.0211	2.4233	2.7045
50	1.2987	1.6759	2.0086	2.4033	2.6778
60	1.2958	1.6706	2.0003	2.3901	2.6603
70	1.2938	1.6669	1.9944	2.3808	2.6479
80	1.2922	1.6641	1.9901	2.3739	2.6387
90	1.2910	1.6620	1.9867	2.3685	2.6316
100	1.2901	1.6602	1.9840	2.3642	2.6259

6 Kritische Grenzen beim Kolmogorov-Smirnov-Anpassungstest

Die Tabelle gibt folgende kritischen Grenzen für den Einstichprobentest an: Beim einseitigen Test: $d^+_{n;1-\alpha}$ maximal mit $P(D^+_n \leq d^+_{n;1-\alpha}) \leq 1-\alpha$; beim zweiseitigen Test: $d_{n;1-\alpha}$ maximal mit $P(D_n \leq d_{n;1-\alpha}) \leq 1-\alpha$.

einseitig: $d^+_{n;1-\alpha}$	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.025$	$\alpha = 0.05$	$\alpha = 0.1$
zweiseitig: $d_{n;1-\alpha}$	$\alpha = 0.01$	$\alpha = 0.02$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$
$n = 1$	0.995	0.990	0.975	0.950	0.900
2	0.929	0.900	0.842	0.776	0.864
3	0.829	0.785	0.708	0.636	0.565
4	0.734	0.689	0.624	0.565	0.493
5	0.669	0.627	0.563	0.509	0.447
6	0.617	0.577	0.519	0.468	0.410
7	0.576	0.538	0.483	0.436	0.381
8	0.542	0.507	0.454	0.410	0.358
9	0.513	0.780	0.430	0.387	0.339
10	0.489	0.457	0.409	0.369	0.323
11	0.468	0.437	0.391	0.352	0.308
12	0.449	0.419	0.375	0.338	0.296
13	0.432	0.404	0.361	0.325	0.285
14	0.418	0.390	0.349	0.314	0.275
15	0.404	0.377	0.338	0.304	0.266
16	0.392	0.366	0.327	0.295	0.258
17	0.381	0.355	0.318	0.286	0.250
18	0.371	0.346	0.309	0.279	0.244
19	0.361	0.337	0.301	0.271	0.237
20	0.352	0.329	0.294	0.265	0.232
21	0.344	0.321	0.287	0.259	0.226
22	0.337	0.314	0.281	0.253	0.221
23	0.330	0.307	0.275	0.247	0.216
24	0.323	0.301	0.269	0.242	0.212
25	0.317	0.295	0.264	0.238	0.208
26	0.311	0.290	0.259	0.233	0.204
27	0.305	0.284	0.254	0.229	0.200
28	0.300	0.279	0.250	0.225	0.197
29	0.295	0.275	0.246	0.221	0.193
30	0.290	0.270	0.242	0.218	0.190
31	0.285	0.266	0.238	0.214	0.187
32	0.281	0.262	0.234	0.211	0.184
33	0.277	0.258	0.231	0.208	0.182
34	0.273	0.254	0.227	0.205	0.179
35	0.269	0.251	0.224	0.202	0.177
Näherung für $n > 40$	$\frac{1.6276}{\sqrt{n}}$	$\frac{1.5174}{\sqrt{n}}$	$\frac{1.3581}{\sqrt{n}}$	$\frac{1.2239}{\sqrt{n}}$	$\frac{1.0730}{\sqrt{n}}$

(Quelle: Bosch: Taschenbuch der Statistik)

7 Übungsaufgaben

Übung 1:

Gegeben seien eine standardnormalverteilte Zufallsvariable X sowie eine Funktion $g : \mathbb{R} \to \mathbb{R}$ mit:

$$g(x) = |x| + 4 \ , \ x \in \mathbb{R}.$$

(a) Bestimmen Sie $P(g(X) \in [a, b])$, $a, b \in \mathbb{R}$, mit $a < b$.

(b) Berechnen Sie $P(g(X) \leq 4.4)$.

(c) Berechnen Sie $E(g(X))$

Übung 2:

X sei eine reellwertige Zufallsvariable, deren Verteilung durch die folgende von einem Parameter $\gamma > 0$ abhängende Dichtefunktion $f_\gamma(x)$ bestimmt ist:

$$f_\gamma(x) \quad = \quad \begin{cases} \frac{1}{\gamma} \cdot (1-x)^{(\frac{1}{\gamma}-1)} & \text{für } 0 < x < 1 \\ \\ 0 & \text{sonst} \end{cases} \quad ,$$

(a) Skizzieren Sie den Verlauf der Dichtefunktion $f_\gamma(x)$ für die Fälle $\gamma = 0.5$ und $\gamma = 2$.

(b) Berechnen Sie die Verteilungsfunktion $F_\gamma(t)$ der Zufallsvariablen X.

(c) Berechnen Sie für beliebiges $n \in \mathbb{N}$ folgende Momente:

$$E(1-X)^n, \ E(X), \ Var(1-X), \ Var(X)$$

(d) Berechnen Sie eine Dichte der Zufallsvariablen $Y := -\ln(1-X)$

Übung 3:

Eine Zufallsvariable X mit $0 < X < 1$ habe die Verteilungsfunktion

$$F_X(x) \quad := \quad \begin{cases} 0 & \text{für } x \leq 0 \\ 1 - e^{-\frac{\lambda x}{1-x}} & \text{für } 0 < x < 1 \\ 1 & \text{für } x \geq 1 \end{cases} \quad ,$$

mit festem Parameter $\lambda > 0$.

(a) Berechnen Sie die Dichte f_X zu X.

(b) Es sei $Y := T(X)$ die Zufallsvariable, die aus X durch die Transformation

$$T : x \mapsto \frac{x}{1-x}$$

hervorgeht. Bestimmen Sie den Wertebereich von Y und geben Sie eine Dichte g_Y zu Y an. Wie heißt die Verteilung der Zufallsvariablen Y?

(c) Bestimmen Sie

$$E\left(\frac{X}{(1-X)^2}\right)$$

Hinweis: Verwenden Sie die Zerlegung $X = X(1-X) + X^2$

Übung 4:

(a) Die Zufallsvariable X sei standardnormalverteilt. Berechnen Sie für $\mu \in \mathbb{R}$ und $\sigma \in \mathbb{R}_{>0}$ die Dichtefunktion der Zufallvariablen Y mit $Y := \exp(\sigma X + \mu)$.

(b) Berechnen Sie den Modalwert und den Median der Zufallsvariablen Y.

(c) Berechnen Sie den Erwartungswert von Y.

(d) Skizzieren Sie die Dichtefunktion von Y für eine geeignete Wahl der Parameter μ, σ und tragen Sie die berechneten Lageparameter im Schaubild ein.

Bemerkung: Die Verteilung von Y heißt Lognormalverteilung $LN(\mu, \sigma)$ und spielt in der Finanzmathematik eine bedeutende Rolle.

Übung 5:

In einer automatischen Überwachungsanlage besteht $\Omega \subset \mathbb{R}^k$ aus Messwert-Kombinationen, die an k Kontrollinstrumenten prinzipiell beobachtbar sind. Nach Experteneinschätzung gilt ein Bereich $S \subset \Omega$ als kritisch ("Störfall"). Eine Modellrechnung liefert eine Wahrscheinlichkeit $P(S) = 2 \cdot 10^{-4}$. Eine weitere Zuverlässigkeitsanalyse liefert die Wahrscheinlichkeiten $P(A|S) = 0.95$ für Alarm, wenn ein Störfall vorliegt (Entdeckungswahrscheinlichkeit), und außerdem $P(A^c|S^c) = 0.99$ für Nichtalarm im unkritischen Bereich.

Für das technische System sind zu berechnen:

(a) die Wahrscheinlichkeit $P(A)$ für Alarm,

(b) die Wahrscheinlichkeit $P(S^c|A)$, dass ein Alarm ein Fehlalarm ist,

(c) die Wahrscheinlichkeit $P(S|A^c)$ für einen unentdeckten Störfall.

Übung 6:

(a) Bei einem Glückspiel werden 200 Kugeln (100 rote und 100 schwarze) gezielt auf zwei Schälchen verteilt, wobei keines der Schälchen leer sein darf. Anschließend wird mit verbundenen Augen zufällig ein Schälchen ausgewählt und hieraus rein zufällig eine Kugel gezogen.

Ist die gezogene Kugel schwarz hat man gewonnen, andernfalls verloren. Wie muss man die Kugeln aufteilen um die maximale Gewinnchance zu haben? Wie groß ist diese?

(b) Von einem Test zur Identifizierung einer ansteckenden Krankeit ist bekannt, dass der Test zu 99% bei infizierten Personen positiv (Infizierung) ausfällt, während er bei 98% der gesunden Personen negativ (keine Infizierung) ist. Ferner ist bekannt, dass 0.1% der Bevölkerung diese Krankheit hat.

Mit welcher Wahrscheinlichkeit ist man tatsächlich infiziert, wenn man ein positives Testergebnis hat?

Übung 7:

(a) Zwei Firmen stellen ein Produkt unter Verwendung von zwei Herstellungsverfahren her. Aus der Gesamtproduktion beider Firmen werde zufällig ein Teil entnommen. Der Zufallsvektor $X = (X_1, X_2, X_3)$ enthält folgende Angaben:

$$X_1 = \begin{cases} 1 & \text{Firma 1} \\ 2 & \text{Firma 2} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{Verfahren 1} \\ 2 & \text{Verfahren 2} \end{cases}$$

$$X_3 = \begin{cases} 0 & \text{Teil gut} \\ 1 & \text{Teil schlecht} \end{cases}$$

Die folgende Tabelle gibt die gemeinsame Wahrscheinlichkeitsverteilung an:

	Firma 1 ($X_1 = 1$)		Firma 2 ($X_1 = 2$)	
	Verfahren 1 ($X_2 = 1$)	Verfahren 2 ($X_2 = 2$)	Verfahren 1 ($X_2 = 1$)	Verfahren 2 ($X_2 = 2$)
Teil gut ($X_3 = 0$)	0.045	0.4	0.225	0.005
Teil schlecht ($X_3 = 1$)	0.005	0.05	0.225	0.045

Welches Verfahren schneidet insgesamt bzw. bei den einzelnen Firmen besser ab? Berechnen Sie zur Beantwortung dieser Frage die folgenden bedingten Wahrscheinlichkeiten:

$$\begin{array}{rcl} P(X_3 = 1 \mid X_2 = 1) & , & P(X_3 = 1 \mid X_2 = 2) \\ P(X_3 = 1 \mid X_1 = 1, X_2 = 1) & , & P(X_3 = 1 \mid X_1 = 1, X_2 = 2) \\ P(X_3 = 1 \mid X_1 = 2, X_2 = 1) & , & P(X_3 = 1 \mid X_1 = 2, X_2 = 2) \end{array}$$

Interpretieren Sie das Ergebnis!

(b) Die folgende Tabelle enthält das Gesamtbruttoeinkommen, sowie die daraus gezahlte Einkommenssteuer der Jahre 1974 und 1978 in den USA, aufgeschlüsselt nach verschiedenen Einkommensklassen:

Jahreseinkommen (pro Person in US$)	Einkommen (in 1000 US$)	gezahlte Steuer (in 1000 US$)
1974		
$< 5\,000$	41 651 643	2 244 467
$5\,000 - 9\,999$	146 400 740	13 646 348
$10\,000 - 14\,999$	192 688 922	21 449 597
$15\,000 - 99\,999$	470 010 790	75 038 230
$\geq 100\,000$	29 427 152	11 311 672
1978		
$< 5\,000$	19 879 622	689 318
$5\,000 - 9\,999$	122 853 315	8 819 461
$10\,000 - 14\,999$	171 858 024	17 155 758
$15\,000 - 99\,999$	865 037 814	137 860 951
$\geq 100\,000$	62 806 159	24 051 698

Modellieren Sie diesen Sachverhalt mit Hilfe eines Wahrscheinlichkeitsraums, wobei Sie als Grundraum Ω die Menge aller 1974 und 1978 verdienter Dollar wählen. Weiterhin bezeichnen $K_1, \ldots, K_5$ die Ereignisse, dass ein zufällig ausgewählter Dollar zu einer der fünf Einkommensklassen K_j gehört, B das Ereignis, dass der Dollar 1974 verdient wurde und A das Ereignis, dass der Dollar als Einkommenssteuer abgeführt wurde. Berechnen Sie die folgenden bedingten Wahrscheinlichkeiten:

$$\begin{array}{rcl} P(A \mid B) & , & P(A \mid B^c) \\ P(A \mid B, K_j) & , & j = 1 \ldots 5 \\ P(A \mid B^c, K_j) & , & j = 1 \ldots 5 \end{array}$$

Interpretieren Sie das Ergebnis!

Übung 8:

Betrachtet wird eine Maschine mit exponentialverteilter Lebensdauer ($T \sim Exp(\lambda)$).

(a) Wie groß ist die Wahrscheinlichkeit für eine Mindestlebensdauer von $a \in \mathbb{R}_+$ Zeiteinheiten? Wie ändert sich diese Wahrscheinlichkeit für wachsende Parameterwerte λ?

(b) Berechnen und interpretieren Sie die bedingte Wahrscheinlichkeit $P([0, t_0 + a] | [t_0, \infty))$ mit $a \geq 0$. Vergleichen Sie Ihr Ergebnis mit dem aus (a).

(c) Wie lautet die Dichte einer auf $[0, a]$ bzw. $[a, \infty]$ gestutzten Exponentialverteilung? Was fällt Ihnen bei der letztgenannten auf?

Übung 9:

Bestimmen Sie für eine auf ein symmetrisches Intervall $[-c, c]$ gestutzte Normalverteilung $N(0; 1)$ die Dichtefunktion φ und die Verteilungsfunktion Φ, und stellen Sie diese graphisch dar.

Übung 10:

Zur grafischen Darstellung des Zusammenhangs zwischen zwei Ereignissen A und B eines Wahrscheinlichkeitsraums $(\Omega, A(\Omega), P)$ erstellen Sie folgende Zeichnung (siehe Abbildung 1):

Die eine Seite eines Quadrats mit der Kantenlänge 1 teilen Sie im Verhältnis $P(A) : 1-P(A)$ und teilen anschließend, entsprechend dieser Aufteilung, das Quadrat in zwei Streifen. An den beiden senkrecht dazu liegenden Seiten tragen Sie von oben die Strecken $P(B \mid A)$ und $P(B \mid \Omega \backslash A)$ ab und teilen anschließend die zu $P(A)$ bzw. $1 - P(A)$ gehörenden Streifen entsprechend.

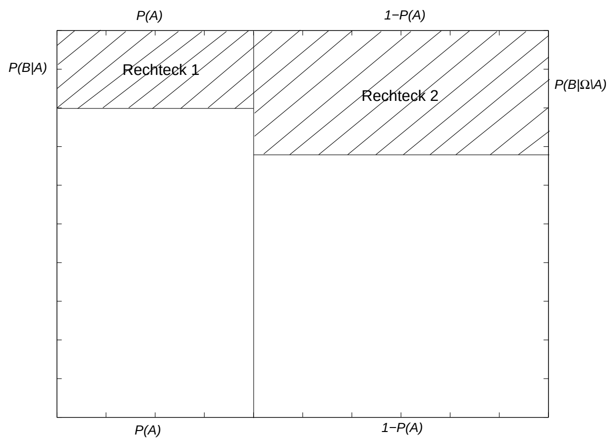

Abbildung 1: Skizze zu Aufgabe 10

(a) Zeigen Sie, dass der Flächeninhalt der schraffierten Fläche gerade $P(B)$ entspricht.

(b) Zeigen Sie, dass die Ereignisse A und B genau dann unabhängig sind, wenn die schraffierte Fläche ein Rechteck bildet.

Übung 11:

(a) Zeigen sie, dass für die Verteilungsfunktion $F(x,y)$ einer zweidimensionalen Zufallsvariablen gilt:

(i) F ist monoton steigend

(ii) F ist in jeder der Variablen von rechts stetig.

(iii)
- $\lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} F(X,Y) = 1$
- $\lim\limits_{x\to-\infty} F(x,y) = 0$ für jedes y
- $\lim\limits_{y\to-\infty} F(x,y) = 0$ für jedes x

(iv) Für alle $x_1 < x_2$ und alle $y_1 < y_2$ gilt:

$$F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \geq 0$$

Bemerkung: Man kann zeigen, dass jede Funktion mit den Eigenschaften (i)-(iv) Verteilungsfunktion einer zweidimensionalen Zufallsvariablen ist.

(b) Überprüfen Sie unter Benutzung von Teil (a), ob die Funktion

(i)

$$F(x,y) = \begin{cases} 1 & \text{für } x \geq 0\ ,\ y \geq 0 \text{ und } x^2+y^2 \geq 4 \\ & \\ 0 & \text{sonst} \end{cases}$$

(ii)

$$F(x,y) \quad = \quad \begin{cases} \min\{x,1\} \cdot \min\{y,1\} \quad \text{für} & x,y \geq 0 \\ & \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion einer zweidimensionalen Zufallsvariablen ist.

Übung 12:

Gegeben sei eine diskrete zweidimensionale Zufallsvariable $Y = (Y_1, Y_2)$ mit der Verteilung

(y_1, y_2)	(-1,0)	(-1,1)	(-1,c)	(2,0)	(2,1)	(2,c)	(c,0)	(c,1)	(c,c)
$P((Y_1, Y_2) = (y_1, y_2))$	1/8	1/4	1/8	1/16	c^2	1/16	1/8	1/8	1/8

(a) Bestimmen Sie c.

(b) Berechnen Sie die Randverteilungen.

(c) Wie groß ist $P(Y \leq (2, c))$?

Übung 13:

Gegeben sei die Funktion

$$\varphi(x, y) := \begin{cases} c(x^2 + y^2) & -1 \leq x, y \leq 1 \\ \\ 0 & \text{sonst} \end{cases}$$

.

(a) Bestimmen Sie $c \in \mathbb{R}$ so, dass φ Dichte einer zweidimensionalen Zufallsvariablen (X, Y) wird.

(b) Berechnen Sie die Randdichten $\varphi_1(x)$ und $\varphi_2(y)$.

(c) Bestimmen Sie die Höhenlinien.

(d) Berechnen Sie $F_{(X,Y)}(0, 0.5)$, $E(X)$ und $E(Y \mid X = \frac{1}{2})$.

Übung 14:

Eine Straßenbahn verkehrt zwischen den Haltestellen G und H. Die Schwarzfahrer unter den Fahrgästen sind zu 60% jugendlich und zu 40% erwachsen. Um diesen das Leben zu erschweren, werden alle Fahrgäste zwischen G und H zweimal kontrolliert, und zwar zuerst von Kontrolleur Nr. 1 und dann von Kontrolleur Nr. 2. Von Kontrolleur Nr.1 entdeckte Schwarzfahrer werden natürlich nicht noch einmal kontrolliert. Kontrolleur Nr. 1 entdeckt 60% der erwachsenen Schwarzfahrer und 40% der jugendlichen Schwarzfahrer. Kontrolleur Nr.2 entdeckt 50% der erwachsenen Schwarzfahrer und 50% der jugendlichen Schwarzfahrer, die vorher noch nicht von Kontrolleur Nr. 1 entdeckt wurden.

(a) Modellieren Sie das Kontrollverfahren durch einen Wahrscheinlichkeitsraum.

(b) Wie groß ist die Wahrscheinlichkeit, dass ein Schwarzfahrer entdeckt wird ?

(c) Wie groß ist die Wahrscheinlichkeit, dass ein entdeckter Schwarzfahrer jugendlich ist ?

(d) Berechnen Sie die Wahrscheinlichkeiten weiterer Kombinationsmöglichkeiten

Übung 15:

In einem Bahnhof steigen drei Reisende in einen Zug ein. Jeder von ihnen setzt sich zufällig in eines von drei Zugabteilen. Die Zufallsvariable N bezeichne die Anzahl der (von ihnen) besetzten Abteile, X_i ($i = 1, 2, 3$) die Anzahl der Reisenden im i-ten Abteil.

(a) Bestimmen Sie die gemeinsame Wahrscheinlichkeitsverteilung von N und X_1 sowie von X_1 und X_2.

(b) Berechnen Sie $E(N)$, $E(X_1)$, $Var(N)$ und $Var(X_1)$.

(c) Sind N und X_1 bzw. X_1 und X_2 unabhängig?

Übung 16:

Die zweidimensionale Zufallsvariable (X, Y) besitze die Dichte

$$f_{(X,Y)}(x, y) = \begin{cases} x + y & \text{für} \quad 0 \leq x, y \leq 1 \\ 0 & \quad\quad \text{sonst} \end{cases} .$$

(a) Bestimmen Sie die Randdichten $f_X(x)$ und $f_Y(y)$.

(b) Berechnen Sie die Verteilungsfunktion von (X, Y).

(c) Sind X und Y unabhängig ?

Übung 17:

Die zweidimensionale Zufallsvariable (X, Y) hat folgende Dichtefunktion:

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{6}{5}x & \text{für} \quad 1 \leq x \leq 2, 0 \leq y \leq x - 1 \\ 0 & \quad\quad \text{sonst} \end{cases}$$

(a) Bestimmen Sie die Randdichten $f_X(x)$ und $f_Y(y)$.

(b) Sind X und Y unabhängig?

(c) Bestimmen Sie Erwartungswert und Varianz von X und Y.

Übung 18:

Eine Münze wird n-mal geworfen. Beim einmaligen Werfen gilt:

$$P(\text{Zahl}){=}p \qquad P(\text{Wappen}){=}1 - p$$

mit $0 < p < 1$. Man geht davon aus, dass die Ergebnisse bei den einzelnen Würfen sich nicht gegenseitig beeinflussen. Beschreiben Sie den Zufallsvorgang "n-maliges Werfen" durch einen Zufallsvektor $X = (X_1 \dots X_n)$ (Verwenden Sie hierbei die Codierung: Wappen=0, Zahl=1). Geben Sie die gemeinsame Wahrscheinlichkeitsverteilung von $X_1 \dots, X_n$ an und bestimmen Sie die Wahrscheinlichkeitsverteilung von $\sum_{i=1}^{n} X_i$. Wie kann $\sum_{i=1}^{n} X_i$ interpretiert werden?

Übung 19:

Überprüfen Sie die folgenden Behauptungen:

(a) Für zwei beliebige diskrete Zufallsvariablen X und Y gilt: Hängt die bedingte Wahrscheinlichkeit von Y unter der Bedingung $X = x$ nicht von x ab, ist also für alle x übereinstimmend, so stimmt die bedingte Wahrscheinlichkeitsverteilung mit der Randverteilung überein (die Umkehrung gilt offensichtlich auch).

(b) Für stetige Zufallsvariablen X und Y gilt die analoge Beziehung zwischen bedingter Dichte und Randdichte.

Übung 20:

Sei (X, Y) eine zweidimensionale Zufallsvariable mit der Dichte

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{1}{4}(x + y + xy) & \text{für} \quad 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie $E(X)$ und $E(Y)$.

(b) Bestimmen Sie die Kovarianzmatrix von (X, Y).

(c) Berechnen Sie den Korrelationskoeffizienten von X und Y.

(d) Sind X und Y unabhängig?

Übung 21:

(a) Bestimmen Sie für die Gleichverteilung über der Fläche $A = \{(x, y)|x^2 + y^2 < 1\} \subset \mathbb{R} \times \mathbb{R}$ die bedingte Verteilung $f(y|x_0)$ für ein beliebiges $x_0 \in (-1, 1)$. Wie sehen die Höhenlinien der bedingten Verteilung in Abhängigkeit von x_0 aus? Sind die beiden Komponenten eines Zufallsvektors mit Gleichverteilung über A unabhängig?

(b) Die Zufallsvariable T gebe einen entsprechend der Gleichverteilung zufällig aus dem Intervall $[0, \alpha]$, $\alpha > 0$, ausgewählten Punkt an. Von diesem Zufallsexperiment werden n unabhängige Wiederholungen gemacht (d.h. wir betrachten n unabhängige Realisationen der Zufallsvariablen T). Die Zufallsvariable X gebe an, wie oft dabei ein Wert aus einem fest vorgegebenen Teilintervall von $[0, \alpha]$ der Länge s gewählt wurde. Wie groß ist die Wahrscheinlichkeit $P(X = k)$?

Übung 22:

Die Anzahl Y der Partikel, die eine Zählkammer in einer bestimmten Zeitspanne erreichen, ist poissonverteilt mit Parameter λ. Von den eintretenden Partikeln wird ein Spannungsstoß ausgelöst, der ein gewisses Vielfaches von Y beträgt. Der zufällig, aber unabhängig von Y erzeugte Multiplikator X hat die Dichtefunktion

$$f(x) = \frac{1}{(1+x)^2} \qquad (x \geq 0) \ .$$

Berechnen Sie die Wahrscheinlichkeit, dass die so entstehende Spannung kleiner als 1 ist.

Übung 23:

Bestimmen Sie die Wahrscheinlichkeitsverteilung, den Erwartungswert und die Varianz der Summe einer $B(m,p)$-verteilten Zufallsvariablen X und einer $B(n,p)$-verteilten Zufallsvariablen Y. (X und Y seien unabhängig)

Übung 24:

(a) Analog zu Aufgabe 18 sei der Ausgang eines Zufallsexperiments durch eine Zufallsvariable X beschrieben. Das Zufallsexperiment wird ohne gegenseitige Beeinflussung n-mal wiederholt. Beschreiben Sie diesen Vorgang durch einen Zufallsvektor $Y = (Y_1 \ldots, Y_n)$ und bestimmen Sie die Verteilungsfunktion von Y, wenn die Verteilungsfunktion von X gegeben ist.

(b) Wie lautet die Verteilungsfunktion, wenn

- (i) X poissonverteilt ist mit Parameter $\lambda > 0$
- (ii) X binomialverteilt ist mit m und p
- (iii) X exponentialverteilt ist mit Parameter $\lambda > 0$
- (iv) X normalverteilt ist mit Mittelwert μ und Varianz σ^2

(c) Wie lautet in den Fällen (b)(i)-(iv) die Verteilungsfunktion von $\sum\limits_{i=1}^{n} Y_i$.

Übung 25:

Veranschaulichen Sie sich den Begriff *Faltung* anhand der beiden unabhängigen diskreten Zufallsvariablen X und Y mit den Wahrscheinlichkeitsverteilungen

$X = x$	1	2	3
$P(X = x)$	0.3	0.2	0.5

$Y = y$	-1	0	1
$P(Y = y)$	0.3	0.4	0.3

Was ändert sich, wenn X und Y stetige Zufallsvariablen sind?

Übung 26:

Die Lebensdauer von einer Anlage sei exponentialverteilt mit Parameter $\lambda_1 > 0$. Bei Ausfall der Anlage wird automatisch und ohne zeitliche Verzögerung ein Notaggregat eingesetzt, dessen Lebensdauer unabhängig von der Lebensdauer der Anlage und ebenfalls exponentialverteilt ist mit einem Parameter λ_2 mit $0 < \lambda_2 \leq \lambda_1$.

(a) Bestimmen Sie die Verteilung der Gesamtlebensdauer des Systems.

(b) Wie lautet die Verteilungsfunktion im Spezialfall $\lambda_2 = \lambda_1$?

(c) Unter der Voraussetzung, dass $\lambda_1 = 2\lambda_2$ ist, ist λ_1 so zu bestimmen, dass eine Gesamtlebensdauer des Systems von mindestens 1000 Zeiteinheiten mit 90% Wahrscheinlichkeit garantiert werden kann.

Übung 27:

Die Lebensdauern T_1 und T_2 zweier elektrischer Bauteile B_1 und B_2 seien exponentialverteilt mit den Parametern $\lambda_1 = \frac{1}{500}$ und $\lambda_2 = \frac{1}{300}$ und unabhängig.

(a) Berechnen Sie die Wahrscheinlichkeit dafür, dass B_1 bzw. B_2 den Zeitpunkt $t_0 = 200$ überlebt, wenn das jeweilige Bauteil zur Zeit $t = 0$ eingesetzt wurde. Wie groß ist die Wahrscheinlichkeit dafür, dass das jeweilige Bauteil nach Erreichen von t_0 noch weitere 200 Stunden arbeitet?

(b) Bestimmen Sie die Lebensdauerverteilung eines aus B_1 und B_2 bestehenden Reihen- bzw. Parallelsystems.

(c) Bestimmen Sie die Wahrscheinlichkeiten dafür, dass das Reihen- bzw. das Parallelsystem den Zeitpunkt $t = 200$ überlebt.

Übung 28:

Die Zufallsvariable X sei gleichverteilt auf dem Intervall $[1, 3]$. $Y = (Y_1, \ldots, Y_n)$ sei eine Stichprobe mit Zurücklegen zu X.

(a) Bestimmen Sie die Verteilung von Y und die Randverteilung von Y_i, $i = 1, \ldots, n$.

(b) Berechnen Sie die Dichte der Spannweite R von Y.

(c) Geben Sie den Erwartungswert von R an.

Übung 29:

(a) Gegeben seien zwei unabhängige diskrete Zufallsvariablen X und Y mit den in der Aufgabe 25 definierten Wahrscheinlichkeitsverteilungen. Bestimmen Sie die Wahrscheinlichkeitsverteilung von $Z = XY$.

(b) Betrachten Sie nun die beiden unabhängigen stetigen Zufallsvariablen X und Y mit den Dichtefunktionen

$$f_X(x) = \begin{cases} \frac{2}{15}x & \text{für} \quad 1 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases},$$

$$f_Y(y) = \begin{cases} \frac{1}{12}y & \text{für} \quad 1 \leq y \leq 5 \\ 0 & \text{sonst} \end{cases}.$$

(i) Bestimmen Sie die Dichtefunktion des Produkts $Z = XY$.

(ii) Überlegen Sie sich, wie Sie die Dichte von $Z = \frac{X}{Y}$ berechnen können.

Übung 30:

(a) Gegeben sei eine diskrete Zufallsvariable $X = (X_1, X_2)$, deren Komponenten unabhängig sind mit den im folgenden angegebenen Randverteilungen. Bestimmen Sie die Wahrscheinlichkeitsverteilung einer Zufallsvariablen $Y = (Y_1, Y_2)$ unter der Annahme, dass $Y_1 = X_1$ und $Y_2 = X_1 + X_2$ gilt.

x_1	-1	0	1
$P(X_1 = x_1)$	0.3	0.4	0.3

x_2	-1	0	1
$P(X_2 = x_2)$	0.3	0.4	0.3

(b) Sei X_1 eine auf dem Intervall $[1, 4]$ gleichverteilte Zufallsvariable und X_2 eine Zufallsvariable mit folgender Dichtefunktion:

$$f_{X_2}(x_2) = \begin{cases} 2 - 0.5 \cdot x_2 & \text{für} \quad 2 \leq x_2 \leq 4 \\ \\ 0 & \text{sonst} \end{cases} .$$

(i) Bestimmen Sie die Dichtefunktion der zweidimensionalen Zufallsvariablen $Y = (Y_1, Y_2)$ unter der Annahme, dass X_1 und X_2 unabhängig sind und $Y_1 = X_1$ sowie $Y_2 = X_1 + X_2$ gilt.

(ii) Bestimmen Sie den Erwartungswert von Y.

Übung 31:

Sei $X = (X_1, X_2)$ eine bivariate Zufallsvariable mit folgender Dichtefunktion:

$$f_X(x_1, x_2) = \begin{cases} 0.1 \cdot x_1 x_2 & \text{für} \quad 1 \leq x_1 \leq 3,\ 2 \leq x_2 \leq 3 \\ \\ 0 & \text{sonst} \end{cases} .$$

Bestimmen Sie die Dichtefunktion der Zufallsvariablen $Y = (Y_1, Y_2)$ mit $Y_1 = X_1$ und $Y_2 = X_1 \cdot X_2$.

Übung 32:

Eine Maschine produziert Stahlstifte mit einer Soll-Länge von $\mu = 110$ mm. Ungenauigkeiten bei der Produktion können leider nicht ausgeschlossen werden, so dass die Längen der produzierten Stifte Realisationen einer Zufallsvariablen X darstellen. Die Streuung der realisierten Stiftlängen ist mit $\sigma = 0.1$ mm bekannt.

Um sich ein Bild von der Güte der Produktionseinheiten zu machen, wird die Wahrscheinlichkeit dafür, dass X höchstens um den Betrag c von μ abweicht, betrachtet. c wird dabei in der Regel in Vielfachen der Standardabweichung σ des Prozesses angegeben, d.h. gesucht ist $P(X \in [\mu - k\sigma, \mu + k\sigma])$.

Bestimmen und vergleichen Sie für $k = 1, 2, 3$ die Wahrscheinlichkeiten für die Fälle, dass

(a) über den Typ der Verteilung von X keine weiteren Kenntnisse verfügbar sind bzw.

(b) X normalverteilt ist mit Mittelwert μ und Varianz σ^2.

Wie erklären Sie sich Ihre Ergebnisse?

Übung 33:

(a) Bei einem Großraumflugzeug ist die Auslastung pro Flug näherungsweise normalverteilt. Im Mittel fliegen 150 Passagiere mit dem Flugzeug, die Auslastung schwankt mit $\sigma = 25$ Passagiere. Mit welcher Anzahl von Fluggästen ist mit einer Wahrscheinlichkeit von 90% mindestens zu rechnen?

(b) Welche Abschätzung können Sie vornehmen, wenn der Typ der Verteilung nicht bekannt ist?

Übung 34:

(a) Erläutern Sie anhand eines idealen Würfels das "schwache Gesetz der großen Zahlen".

(b) Verdeutlichen Sie sich die wesentlichen Aussagen des zentralen Grenzwertsatzes.

(c) Eine Vertriebsgesellschaft besitzt in einer Großstadt 200 Zigarettenautomaten. Jeder Automat hat (unabhängig von den anderen) mit der Wahrscheinlichkeit $\frac{1}{20}$ pro Woche eine Störung. Für die Entscheidung über die Größe eines ständigen Reparaturtrupps sei die Wahrscheinlichkeit dafür von Interesse, dass in einer Woche die Anzahl X der defekten Automaten zwischen 5 und 15 liegt, von Interesse. Diese Wahrscheinlichkeit (der exakte Wert beträgt übrigens 0.9292) soll

   (i) mittels der Poissonverteilung approximiert werden,

   (ii) über die Ungleichung von Tschebyscheff nach unten abgeschätzt werden und

   (iii) mit Hilfe des zentralen Grenzwertsatzes approximativ berechnet werden.

Übung 35:

Ein Unternehmen stellt Taschenlampenbatterien her. Für ein Marketingprojekt soll die Lebensdauer dieser Batterien mit einer Stichprobe untersucht werden, wobei eine Exponentialverteilung unterstellt wird. Geben Sie zu diesem Beispiel die drei Grundannahmen der schließenden Statistik an und erläutern Sie die drei Grundaufgaben. Geben Sie zu jeder der Grundaufgaben eine konkrete Fragestellung für das Unternehmen an, die zu dieser Aufgabe führt, und wie eine Entscheidungsfunktion dazu aussieht. Geben Sie jeweils eine sinnvoll erscheinende Entscheidungsfunktion an.

Übung 36:

Ein Briefmarkensammler weiß, dass von einem bestimmten Ersttagsbrief eine gewisse Auflagenhöhe existiert. Von k dieser Ersttagsbriefe, die wie gewöhnlich mit 1 beginnend durchnumeriert wurden, kennt er per Zufall die Nummern $n_1, \ldots, n_k$. Aufgrund dieser Information möchte er einen Schätzwert für die Gesamtauflage N des Ersttagsbriefes erstellen. Begründen Sie, wie die im folgenden vorgeschlagenen Schätzer zustande gekommen sind:

(a) $N_1 = \max_i n_i$

(b) $N_2 = \frac{k+1}{k} \max_i n_i$

(c) $N_3 = \max_i n_i + \min_i n_i - 1$

(d) $N_4 = 2 \frac{1}{k} \sum_{i=1}^{k} n_i$

(e) $N_5 = 2 \frac{1}{k} \sum_{i=1}^{k} n_i - 1$

Übung 37:

(a) Eine Zufallsvariable sei im Bereich $[0, a]$ gleichverteilt, a ist nicht bekannt und soll mit Hilfe einer Stichprobe mit Zurücklegen geschätzt werden. Überlegen Sie, was die für diese Aufgabe wesentliche Information der Stichprobenwerte $x_1, \ldots, x_n$ ist. Überprüfen Sie, ob diese Stichprobenfunktion suffizient ist.

(b) Eine Zufallsvariable X habe die Dichtefunktion

$$f(x) = \begin{cases} 0 & \text{für} \quad x < a \\ \lambda e^{-\lambda(x-a)} & \text{für} \quad x \geq a \end{cases}$$

mit $a > 0$ und bekanntem λ. Lösen Sie die zu (a) analoge Aufgabe.

Übung 38:

Bestimmen Sie mit Hilfe des Faktorisierungstheorems von Neyman suffiziente Statistiken für

(a) die Klasse der Exponentialverteilungen

(b) die Klasse der Gleichverteilungen auf dem Intervall $[-\vartheta, \vartheta]$ $(\vartheta > 0)$.

(c) die Klasse der Binomialverteilungen $B(m, p)$

Übung 39:

Zeigen Sie, dass z.B. die diskrete Verteilung $B(m, p)$ eine Exponentialfamilie bildet. Wie lauten die daraus resultierenden suffizienten Statistiken?

Übung 40:

(a) Der Torschützenkönig der Fußball-WM trifft an der Torwand mit Wahrscheinlichkeit p. Mit einer Stichprobe vom Umfang 2 soll p geschätzt werden. Es wird folgende Schätzfunktion vorgeschlagen:

Bei 0 Treffern lautet der Schätzwert $\hat{p} = 0$
Bei 1 Treffer lautet der Schätzwert $\hat{p} = \frac{1}{2}$
Bei 2 Treffern lautet der Schätzwert $\hat{p} = 1$

Wie lautet der Erwartungswert bei diesem Schätzverfahren? Versuchen Sie eine Schätzfunktion ausgehend von der Trefferzahl zu bestimmen, die ebenfalls diesen Erwartungswert besitzt, aber verschieden von der hier vorgeschlagenen ist.

(b) Die Zufallsvariable X sei Weibull-verteilt mit Parametern α ($\alpha > 0$) und β, d.h. ihre Dichtefunktion lautet:

$$f_X(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} & \text{für} \quad x > 0 \\ 0 & \text{sonst} \end{cases} .$$

(i) Zeichnen Sie die Dichtefunktion für einige Parameterkombinationen (α, β). Welchen Einfluß haben die Parameter auf die Gestalt der Dichtefunktion?

(ii) Der Wert von β sei nun mit $\beta = 3$ bekannt, α soll mit Hilfe einer einfachen Stichprobe vom Umfang n geschätzt werden. Bestimmen Sie eine suffiziente Statistik für α. Ist die Statistik vollständig?

Übung 41:

Gegeben seien die folgenden drei anhand einer zufälligen Stichprobe mit Zurücklegen vom Umfang n gebildeten Schätzfunktionen für das arithmetische Mittel μ einer Grundgesamtheit, wobei $E(X_i) = \mu$ und $Var(X_i) = \sigma^2 < \infty$:

- $T_1(x_1, \ldots, x_n) = \frac{1}{n} \sum\limits_{i=1}^{n} x_i$
- $T_2(x_1, \ldots, x_n) = x_j$ für ein festes $j \in \{1, \ldots, n\}$
- $T_3(x_1, \ldots, x_n) = \frac{2}{3}(\frac{1}{n-1} \sum\limits_{i=1}^{n-1} x_i) + \frac{1}{3}x_n$

(a) Welche der drei Schätzfunktionen sind erwartungstreu, d.h. liefern „im Mittel“ das arithmetische Mittel μ?

(b) Von welchem dieser Schätzer würden Sie - ohne zu rechnen - annehmen, dass er die größte Varianz besitzt? Warum?

(c) Berechnen Sie die zugehörigen Varianzen, und ordnen Sie diese der Größe nach.

Zur Schätzung des arithmetischen Mittels μ seien nun alle linearen Schätzer der Form $\sum_{i=1}^{n} \alpha_i x_i$, $\alpha_i \in \mathbb{R}$, zugelassen.

(d) Zeigen Sie, dass die drei oben angegebenen Schätzfunktionen linear sind, indem Sie die jeweiligen Gewichte α_i bestimmen.

(e) Welche Bedingung an die Gewichte α_i garantiert, dass die entsprechenden Schätzer „im Mittel“ das arithmetische Mittel μ liefern?

(f) Bestimmen Sie unter allen linearen Schätzfunktionen, die die in (e) definierte Bedingung erfüllen, diejenige mit minimaler Varianz.

Übung 42:

Zeigen Sie, dass die in Übung 36 untersuchte Schätzfunktion

$$N_5 = 2\frac{1}{k}\sum_{i=1}^{k} n_i - 1$$

erwartungstreu ist. Dabei gehe man davon aus, dass es sich in Aufgabe 36 um eine Stichprobe mit Zurücklegen handelt, bei der jeder Ersttagsbrief dieselbe Chance hat, gezogen zu werden.

Übung 43:

Die Dichtefunktion der Gammaverteilung lautet ($\alpha, \lambda > 0$)

$$f_Y(y) = \begin{cases} \frac{1}{\Gamma(\alpha)}\lambda^\alpha y^{\alpha-1} e^{-\lambda y} & \text{für} \quad y > 0 \\ & \\ 0 & \text{sonst} \end{cases},$$

dabei ist $\Gamma(\alpha)$ der Wert der sogenannten Gammafunktion an der Stelle α (eine explizite Darstellung dieser Funktion wird im weiteren nicht benötigt).

(a) Skizzieren Sie die Dichtefunktion für einige Parameterkombinationen (α, λ). Welchen Einfluß haben die Parameter auf die Gestalt der Dichtefunktion?

(b) Bestimmen Sie für eine einfache Stichprobe mit Zurücklegen eine suffiziente und vollständige Statistik für das Parameterpaar (α, λ).

(c) Der Erwartungswert der Gammverteilung lautet $E(Y) = \frac{\alpha}{\lambda}$. Es sei bekannt, dass $\lambda = 1$ gilt. Sei X eine einfache Stichprobe mit Zurücklegen, die zum Schätzen des Parameters α gezogen wurde. Untersuchen Sie, ob $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ eine gleichmäßig beste erwartungstreue Schätzfunktion für α ist.

Übung 44:

Verdeutlichen Sie sich an folgendem Beispiel das Prinzip der *Maximum-Likelihood-Schätzung*: In einem Teich soll die Anzahl N der Fische geschätzt werden. Dazu werden M Fische gefangen und markiert. Anschließend werden sie wieder freigelassen. Nach einigen Tagen werden erneut Fische gefangen. Von den n Fischen des Fangs seien x markiert.

Für welches N hat die beobachtete Stichprobe die größtmögliche Wahrscheinlichkeit?

Übung 45:

Es wurden 200 Gruppen von jeweils 10 Werkstücken auf Fehler untersucht. Die Zufallsvariable X gebe jeweils die Anzahl der defekten Werkstücke an:

Anzahl defekter Werkstücke	0	1	2	3	≥ 4
beobachtete Häufigkeit	133	52	12	3	0

Aus Erfahrung weiß man, dass X binomialverteilt ist. Bestimmen Sie die Maximum-Likelihood-Schätzfunktion (ML-Schätzer) für den Parameter p dieser Verteilung. Berechnen Sie anschließend den ML-Schätzwert für die gegebenen Stichprobenergebnisse. Wie lauten unter dem geschätzten Binomialmodell die Wahrscheinlichkeiten für die Ereignisse entsprechend der Tabelle?

Übung 46:

In einer automatischen Abfüllanlage wird Zucker in 1-Pfund-Tüten abgefüllt, wobei es zu Abweichungen vom Normgewicht kommen kann. Der Hersteller der Anlage gibt an, dass bei Einstellung einer bestimmten Soll-Füllmenge θ Abweichungen um maximal 5 g in beide Richtungen möglich seien. Weiterhin sollen alle Abweichungen gleich wahrscheinlich sein.

Zur Überprüfung der Einstellung der Anlage wird eine Stichprobe vom Umfang 20 gezogen. Folgende Gewichtswerte werden gemessen:

502	501	500	498	497	499	500	505	496	495,5
499	502	501	497	496	496	505	502	500	496

Bestimmen Sie auf Basis dieser Werte einen ML-Schätzwert für die tatsächlich eingestellte Abfüllmenge θ.

Übung 47:

Eine Tankstelle veranschlagt für einen Ölwechsel mindestens α Minuten. Die tatsächlich benötigte Zeit X $(X \geq \alpha)$ ist von Kunde zu Kunde verschieden, wobei man davon ausgehen kann, dass $X - \alpha$ annähernd exponentialverteilt $(\lambda = 1)$ ist. Zur Schätzung von α wurden folgende Arbeitszeiten notiert:

4.2	3.1	3.6	4.5	5.1	7.6	4.4	3.5	3.8
3.9	4.1	4.3	4.4	3.7	3.5	3.8	3.9	

Bestimmen Sie

(a) einen Schätzwert für α mit Hilfe der Maximum-Likelihood-Schätzmethode.

(b) einen Schätzwert für die Zeit, die im Mittel für einen Ölwechsel benötigt wird.

Übung 48:

In einer Zufallsstichprobe von 250 Studenten einer Universität waren 22 Linkshänder.

(a) Geben Sie einen Bereich an, der mit 95% Wahrscheinlichkeit den Anteil p der Linkshänder unter den Studenten dieser Universität enthält.

(b) Berechnen Sie ein angenähertes 95%-Konfidenzintervall für p unter Verwendung des zentralen Grenzwertsatzes. Wählen Sie dabei die drei folgenden verschiedenen Vorgehensweisen:

  (i) Schätzen Sie die die unbekannte Standardabweichung durch die Stichprobenstandardabweichung.

  (ii) Schätzen Sie die unbekannte Standardabweichung durch die korrigierte Stichprobenstandardabweichung.

  (iii) Verwenden Sie den exakten (aber unbekannten) Wert für die Standardabweichung.

Übung 49:

Das Gewicht von Wäscheklammern sei als näherungsweise normalverteilt angenommen. Folgende Daten wurden in einer Stichprobe gemessen (in $[g]$):

9.3	8.5	9.8	10.1	7.2	11.4	12.5	10.8
11.3	13.6	10.3	8.2	12.8	15.2	9.6	10.7

(a) Bestimmen Sie ein 95%- und ein 99%-Konfidenzintervall für den Mittelwert und die Varianz des Gewichts der Klammern.

(b) Von der Firma, die die Wäscheklammern produziert, erfahren Sie, dass das durchschnittliche Klammergewicht 10.5 g beträgt. Bestimmen Sie nun ein 95%-Konfidenzintervall für die Varianz.

(c) Bestimmen Sie ein 95%-Konfidenzintervall für den Mittelwert unter der Annahme, dass die Varianz des Klammergewichts mit $\sigma^2 = 5\ g^2$ bekannt ist.

Übung 50:

Y sei Poisson-verteilt mit Parameter λ. Zu testen sei die Hypothese $H_0 :\ \lambda = \lambda_0$ gegen $H_1 :\ \lambda = \lambda_1$ mit $\lambda_1 > \lambda_0$. Konstruieren Sie einen einfachen besten Test zum Niveau α.

Übung 51:

Um beurteilen zu können, wie sich eine 10%-ige Preiserhöhung bei einem Artikel auf dessen Absatz auswirkt, werden Verkaufsstellen ausgewählt, in denen probeweise eine Preiserhöhung durchgeführt wird. Die Preiserhöhung soll dann generell durchgeführt werden, wenn mit einer durchschnittlichen Absatzminderung von höchstens 8% zu rechnen ist. Es werden nun zufällig n Verkaufsstellen ausgewählt und die sich dort ergebenden Absatzveränderungen bestimmt.

(a) Erläutern Sie anhand der Situationsbeschreibung kurz, was unter einem *Test* zu verstehen ist. Überlegen Sie dabei insbesondere, welches Ziel Sie beim Testen verfolgen, welche Informationen Ihnen gegeben sind und mit Hilfe welcher Schritte Sie basierend auf den Ihnen zur Verfügung stehenden Informationen Ihr Ziel erreichen können.

(b) Formulieren Sie ein geeignetes statistisches Entscheidungsproblem (Modell und Hypothesen).

(c) Welche Fehlentscheidungen können in diesem Beispiel auftreten?

(d) Welcher Zusammenhang besteht zwischen den Entscheidungen und der Wahl der Nullhypothese?

(e) Wie beeinflußt die Wahl der Nullhypothese die Entscheidung des Testverfahrens?

Übung 52:

Ein Unternehmen erhält von einem Lieferanten den Hinweis, dass eine gelieferte Warenpartie von 10 000 Teilen möglicherweise nicht den gewöhnlichen Ausschußanteil von 5%, sondern aufgrund von Produktionsschwierigkeiten einen erhöhten Anteil von 15% besitzt. Zur Überprüfung dieses Hinweises wird eine Stichprobe vom Umfang $n = 30$ entnommen. Mit einem gleichmäßig besten Test zu einem Niveau von höchstens 0.1 soll „statistisch“ gesichert werden, dass die Partie wie normal einen Ausschussanteil von 5% hat. Formulieren Sie den Test.

Übung 53:

Eine Firma stellt Leuchtstoffröhren her. Es wurde ein neues Produktionsverfahren eingeführt, aufgrund dessen sich die mittlere Lebensdauer der Röhren (bisher 1500 Stunden) erhöhen soll. Man möchte prüfen, ob dies tatsächlich der Fall ist und entnimmt daher eine Stichprobe vom Umfang 10 aus der laufenden Produktion.

(a) Formulieren Sie die Nullhypothese, und geben Sie einen gleichmäßig besten Test zum Niveau $\alpha = 0.05$ an, wenn angenommen wird, dass die Lebensdauer der Röhren normalverteilt ist und die Standardabweichung $\sigma = 100\ h$ beträgt.

(b) Sei $(1430, 1480, 1520, 1500, 1550, 1460, 1530, 1540, 1600, 1510)$ das Ergebnis der Stichprobe. Welche Entscheidung trifft das Unternehmen?

Übung 54:

Eine Firma hat über einen Zeitraum von 45 Wochen den Absatz eines ihrer Produkte ermittelt. Dabei ergab sich ein Stichprobenmittel $\bar{x} = 447\,000$ Stück/Woche. Die Firma weiß weiterhin aus langjähriger Erfahrung, daß der Absatz dieses Produktes normalverteilt ist mit dem Mittelwert μ und der Standardabweichung $\sigma = 20\,000$.

(a) Die Firma möchte nun testen, ob die Stichprobe der Annahme widerspricht, dass der durchschnittliche Absatz pro Woche 450 000 Stück beträgt.

- Formulieren Sie die Null- und die Gegenhypothese.
- Geben Sie einen Test zum Signifikanzniveau 0.05 an.
- Welche Entscheidung wird die Firma treffen?

(b) Die Firma hat nun eine Anzeigenkampagne gestartet. Bei einer anschließenden Untersuchung über 16 Wochen werden ein Stichprobenmittel von $\bar{x} = 460\,000$ und eine korrigierte Stichprobenstandardabweichung von $s^* = 21\,000$ festgestellt. Geben Sie einen Test zum Signifikanzniveau 0.05 an, und ermitteln Sie, ob die Anzeigenkampagne erfolgreich war. Dabei werde eine unbekannte Varianz zugrunde gelegt.

(c) Was beschreibt die Gütefunktion eines Tests allgemein? Skizzieren Sie die Gütefunktion für einen einseitigen Test beispielhaft anhand der Daten aus Aufgabenteil (a). Wie würde Ihrer Meinung nach eine ideale Gütefunktion aussehen?

Übung 55:

Zehn Hohlkarabiner einer bestimmten Marke wurden der Produktion zufällig entnommen und einem Zerreißversuch unterzogen, bei dem die Belastung eines Karabiners jeweils so lange kontinuierlich erhöht wurde, bis er brach. Der Bruch trat bei folgenden Werten (in $[kp]$) auf:

$$2100, 2130, 2150, 2170, 2210, 2070, 2230, 2150, 2230, 2200$$

(Daten aus "Deutscher Alpenverein", 1, München 1979, Februar).

(a) Testen Sie unter der Voraussetzung, dass die Karabiner bzgl. ihrer Bruchlast X einer Normalverteilung genügen, die Hypothese

$$H_0: \ \sigma^2 = 1600 \qquad \text{gegen} \qquad H_1: \ \sigma^2 \neq 1600,$$

wobei σ^2 die Varianz der Bruchlast in der Gesamtheit aller produzierten Karabiner des untersuchten Typs bedeute, zum Niveau $\alpha = 0.05$.

(b) Es besteht die Vermutung, dass die tatsächliche Varianz größer als $\sigma^2 = 1600$ ist. Führen Sie einen geeigneten Test durch.

(c) Führen Sie den Test aus (a) durch, falls Sie den Erwartungswert mit $\mu = 2100\ kp$ kennen.

Übung 56:

(a) In der englischsprachigen Literatur sowie bei der Verwendung von Software-Paketen werden Sie im Zusammenhang mit Tests oftmals auf den so genannten "p-value" stossen. Erklären Sie, was man unter diesem Begriff versteht. Geben Sie an, wie sich Ihre Vorgehensweise beim Testen ändert, wenn Sie über den p-Wert argumentieren wollen.

(b) Geben Sie bei den Aufgaben 53-55 jeweils den p-Wert an.

Übung 57:

Im Rahmen einer medizinischen Untersuchung wird das Gewicht von 15-jährigen Jungen bestimmt. Es ergaben sich folgende 20 Werte:

49.1	55.0	44.9	53.8	60.4	51.6	53.2	41.2	58.3	50.4
56.1	56.5	47.8	43.6	60.5	47.3	59.7	55.2	57.1	54.5

(a) Aus medizinischer Sicht interessant ist die Frage, ob das Gewicht 15-jähriger Jungen approximativ als normalverteilt angenommen werden kann. Überlegen Sie sich, wie man eine solche Vermutung testen könnte. Gehen Sie dabei insbesondere auf die Punkte

- Hypothesenwahl
- Formulierung einer Teststatistik
- Annahme- bzw. Ablehnungsbereich

ein.

(b) Testen Sie die Hypothese, dass das Gewicht 15-jähriger Jungen normalverteilt ist mit $\mu = 50$ und $\sigma = 5$. Verwenden Sie dazu den Kolmogorov-Smirnov-Anpassungstest.

(c) Könnte man mit Hilfe des Kolmogorov-Smirnov-Anpassungstests auch überprüfen, ob die Gewichtsverteilung 15-jähriger Jungen einer Normalverteilung $N(50, \sigma^2)$ mit unbekanntem σ^2 entspricht?