Wahrscheinlichkeitstheorie und Statistik, Prüfung, mit Lösungen

M.Gruber

16.Juli 2009, 13:30–15:00, R 1.049 (59)

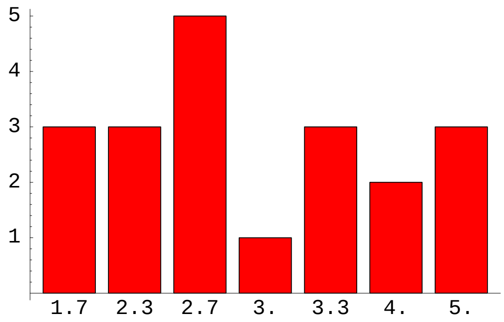

Abbildung 1: Notenstatistik

1. (10 Punkte) *(Bedingte) Wahrscheinlichkeit*

   Für die "Neue Grippe" gibt es einen Labortest. In 10% der Fälle liefert er fälschlicherweise das Resultat "negativ", obwohl die Krankheit vorliegt. In 30% der Fälle liefert er fälschlicherweise das Resultat "positiv", obwohl die Krankheit nicht vorliegt. Für die Aufgabe nehmen wir an, dass 5% der Bevölkerung an der "Neue Grippe" erkrankt sind.

   (a) Wie groß ist die Wahrscheinlichkeit, dass der Labortest ein korrektes Ergebnis liefert?

   (b) Wie groß ist die Wahrscheinlichkeit, dass eine Person tatsächlich an der "Neuen Grippe" erkrankt ist, wenn der Labortest das Resultat "positiv" liefert?

   (c) Ein "alternativer Test" liefere stets das Resultat "negativ".
   Wie groß ist die Wahrscheinlichkeit, dass der "alternative Test" ein korrektes Ergebnis liefert?

   Lösung $A :=$ "Neue Grippe" liegt vor, $Pos :=$ Test liefert "positiv", $Neg :=$ Test liefert "negativ".

   Den Angaben entnimmt man: $P(Neg \mid A) = 0.1$, $P(Pos \mid \neg A) = 0.3$, $P(A) = 0.05$.

   Folgerungen: $P(Pos \mid A) = 1 - P(Neg \mid A) = 0.9$, $P(Neg \mid \neg A) = 1 - P(Pos \mid \neg A) = 0.7$. $P(\neg A) = 1 - P(A) = 0.95$.

   Antworten:

   (a) $P(\text{Test korrekt}) = P(Pos \wedge A) + P(Neg \wedge \neg A) = P(Pos \mid A)P(A) + P(Neg \mid \neg A)P(\neg A) = 0.9 \cdot 0.05 + 0.7 \cdot 0.95 = 0.71$.

   (b) $P(A \mid Pos) = \dfrac{P(Pos \mid A)P(A)}{P(Pos \mid A)P(A) + P(Pos \mid \neg A)P(\neg A)} = \dfrac{0.9 \cdot 0.05}{0.9 \cdot 0.05 + 0.3 \cdot 0.95} = 0.136364$.

(c) $P(\text{Alternativtest korrekt}) = P(\neg A) = 0.95$.

2. (10 Punkte) *Gemeinsame Verteilung*

   Seien $X, Y \sim \mathcal{U}[0,1]$ unabhängige Zufallsvariablen.

   Gesucht ist die Dichte von $X - Y$.

   Hinweise: Eine Unterscheidung der Fälle $-1 \leq z \leq 0$ und $0 < z \leq 1$ für $X - Y \leq z$ kann hilfreich sein.

   Lösung Sei $F(z) = P(X - Y \leq z)$.

   $-1 \leq z \leq 0$: $F(z) = P(Y \geq X - z) = \int_0^{1+z} \int_{x-z}^1 dy\, dx = (1+z)^2/2 \Rightarrow F'(z) = 1 + z$.

   $0 < z \leq 1$: $F(z) = P(Y \geq X - z) = \int_0^z \int_0^1 dy\, dx + \int_z^1 \int_{x-z}^1 dy\, dx = 1 - (1-z)^2/2 \Rightarrow F'(z) = 1 - z$.

   Somit ist die gesuchte Dichte $f(z) = (1 - |z|)[-1 < z < 1]$.

3. (10 Punkte) *Transformation einer Zufallsvariablen*

   Die Zufallsvariable X habe die Dichte $f_X(x) = 3x^2[0 < x < 1]$.

   Welche Dichte hat $1/X$?

   Lösung

   $Y := 1/X$. $f_Y(y) = f_X(1/y) \cdot |-1/y^2| = 3 \cdot (1/y^2) \cdot [0 < 1/y < 1] \cdot (1/y^2) = (3/y^4) \cdot [1 < y < \infty]$.

4. (10 Punkte) *Korrelation*

   Seien $X, Y \sim \mathcal{U}[0,1]$ unabhängige Zufallsvariablen.

   Berechnen Sie die Korrelation $\rho(X, X + \frac{1}{2}Y)$.

   Lösung

   $\operatorname{Var} Y = \operatorname{Var} X = \mathbf{E}X^2 - (\mathbf{E}X)^2 = \int_0^1 x^2\, dx - 1/4 = 1/12 \Rightarrow \sigma(X) = \sqrt{1/12}$.

   $\operatorname{Var}(X + (1/2) \cdot Y) = \operatorname{Var} X + (1/4) \cdot \operatorname{Var} Y = 5/48 \Rightarrow \sigma(X + (1(2) \cdot Y) = (1/2)\sqrt{5/12}$.

   $\operatorname{Cov}(X, Y) = 0$.

   $\operatorname{Cov}(X, X + (1/2) \cdot Y) = \operatorname{Var} X + (1/2) \cdot \operatorname{Cov}(X, Y) = \operatorname{Var} X$.

   $\rho(X, X + (1/2) \cdot Y) = \operatorname{Cov}(X, X + (1/2) \cdot Y)/(\sigma(X) \cdot \sigma(X + (1/2) \cdot Y)) = \sigma(X)/\sigma(X + (1/2) \cdot Y) = 2/\sqrt{5}$.

5. (10 Punkte) *Normalverteilung*

   Die Zufallsvariable X sei normalverteilt mit Mittelwert 16 und Varianz 25.

   Berechnen Sie die Wahrscheinlichkeit $P(20 < X - 1 \leq 30)$.

   Lösung

   $P(20 < X - 1 \leq 30) = P(5 < X - 16 \leq 15) = P(1 < (X - 16)/5 \leq 3) = \Phi(3) - \Phi(1) = 0.9987 - 0.8413 = 0.1574$.

6. (10 Punkte) *Maximum-Likelihood-Schätzung*

   Während der Vorlesung kommen dauernd Studenten zu spät. Der Dozent beobachtet die folgenden drei Zeitabstände (in Minuten): 5,8,5. Er nimmt an, dass die Zeitabstände exponentialverteilt sind mit Dichte $f(t \mid \alpha) = \alpha e^{-\alpha t}[0 < t]$ und schätzt α mit der Maximum-Likelihood-Methode.

   Welchen Schätzwert $\hat{\alpha}$ erhält er?

   Lösung Maximum-Likelihood-Funktion: $\psi(\alpha) = \alpha^3 e^{-\alpha(5+8+5)} = \alpha^3 e^{-18\alpha}$.

$\psi'(\alpha) = 0 \Leftrightarrow (\ln\psi)'(\alpha) = 0.$

$\ln\psi(\alpha) = 3\ln\alpha - 18\alpha.$

$(\ln\psi)'(\alpha) = 3/\alpha - 18 = 0 \Rightarrow \alpha = 1/6.$

$\hat{\alpha} = 1/6.$

7. (10 Punkte) *Momenterzeugende Funktion*

   Die Zufallsvariable X habe die momenterzeugende Funktion $\varphi(u) = \mathbf{E}e^{uX} = -\alpha/(u-\alpha)$.

   Welchen Wert haben $\mathbf{E}X$ (1.Moment) und $\mathbf{E}X^2$ (2.Moment) ?

   Lösung $\varphi'(u) = \alpha/(u-\alpha)^2$. Damit ist $\mathbf{E}X = \varphi'(0) = 1/\alpha$.

   $\varphi''(u) = -2\alpha(u-\alpha)/(u-\alpha)^4 = -2\alpha/(u-\alpha)^3$. Damit ist $\mathbf{E}X^2 = \varphi''(0) = 2/\alpha^2$.

8. (10 Punkte) *Konfidenzintervall*

   Der Milchsäuregehalt einer Käsesorte sei normalverteilt mit unbekanntem Mittelwert μ und unbekannter Varianz σ^2.

   Es werden sechs Stichproben durchgeführt mit den Ergebnissen $0.9, 1.1, 0.9, 0.9, 1.1, 1.1$.

   Geben Sie das Konfidenzintervall für μ mit Konfidenzwahrsscheinlichkeit 0.9 an.

   Lösung $\overline{x^2} = 1.01$, $\overline{x}^2 = 1.0$. $\hat{\sigma^2} = \overline{x^2} - \overline{x}^2 = 0.01$.

   Das 0.95-Quantil der t_5-Verteilung ist $c = 2.015$.

   Das Konfidenzintervall ist $[\overline{x} - c\sqrt{(1/5)\hat{\sigma^2}}, \overline{x} + c\sqrt{(1/5)\hat{\sigma^2}}] = [0.909886, 1.09011]$.

9. (10 Punkte) *Statistischer Test*

   Ein Medikament hilft angeblich in mindestens 90% aller Fälle. Diese Aussage soll durch eine Stichprobe vom Umfang 121 möglichst nachgewiesen werden (Signifikanzniveau $\alpha = 0.05$).

   (a) Formulieren Sie den Test (Hypothesen H_0, H_1, Testgröße, Verteilung der Testgröße, Annahmebereich).

   (b) Bei der Stichprobe wird in 115 Fällen ein Heilerfolg festgestellt.
   Wie lautet die Entscheidung?
   Wie zuverlässig ist die Entscheidung?

   Lösung

   (a) $H_0 : p \leq 0.9, H_1 : p > 0.9$.
   Testgröße $T = \frac{1}{121}\sum_{1\leq i\leq 121} X_i$ mit $X_i = 1$ falls bei der i-ten Stichprobe eine erfolgreiche Behandlung festgestellt wird, andernfalls $X_i = 0$.
   Die Testgröße ist annähernd normalverteilt.
   Annahmebereich: $[0, 0.9 + 1/242 + 1.645\sqrt{0.9\cdot 0.1/121}] = [0, 0.948996]$.

   (b) $115/121 = 0.950413$. Die Nullhypothese wird abgelehnt.
   Es ist damit statistisch nachgewiesen, dass das Medikament in mindestens 90% der Fälle hilft.
   Die Entscheidung ist im ungünstigsten Fall mit Wahrscheinlichkeit 0.05 falsch.

Werte der Standardnormalverteilung $\Phi(z) = P(X \leq z)$ für $z \geq 0$

z	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	0.5080	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	0.7704	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	0.8599	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	0.9066	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	0.9808	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	0.9868	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	0.9959	0.9960	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

Obere Perzentile der t-Verteilung mit df Freiheitsgraden

df	90%	95%	97.5%	99%	99.5%	99.9%
1	3.078	6.314	12.706	31.821	63.657	318.313
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.782
8	1.397	1.860	2.306	2.896	3.355	4.499
9	1.383	1.833	2.262	2.821	3.250	4.296
10	1.372	1.812	2.228	2.764	3.169	4.143
11	1.363	1.796	2.201	2.718	3.106	4.024
12	1.356	1.782	2.179	2.681	3.055	3.929
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
70	1.294	1.667	1.994	2.381	2.648	3.211
80	1.292	1.664	1.990	2.374	2.639	3.195
90	1.291	1.662	1.987	2.368	2.632	3.183
100	1.290	1.660	1.984	2.364	2.626	3.174
∞	1.282	1.645	1.960	2.326	2.576	3.090