Tabelle 3: Normalverteilung

$\phi(z) = 0,\ldots$

$\phi(-z) = 1 - \phi(z)$

z	0	1	2	3	4	5	6	7	8	9
0,0	5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0,1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
0,2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
0,3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0,4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0,5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0,6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
0,7	7580	7611	7642	7673	7704	7734	7764	7794	7823	7852
0,8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0,9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1,0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8621
1,1	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830
1,2	8849	8869	8888	8907	8925	8944	8962	8980	8997	9015
1,3	9032	9049	9066	9082	9099	9115	9131	9147	9162	9177
1,4	9192	9207	9222	9236	9251	9265	9279	9292	9306	9319
1,5	9332	9345	9357	9370	9382	9394	9406	9418	9429	9441
1,6	9452	9463	9474	9484	9495	9505	9515	9525	9535	9545
1,7	9554	9564	9573	9582	9591	9599	9608	9616	9625	9633
1,8	9641	9649	9656	9664	9671	9678	9686	9693	9699	9706
1,9	9713	9719	9726	9732	9738	9744	9750	9756	9761	9767
2,0	9772	9778	9783	9788	9793	9798	9803	9808	9812	9817
2,1	9821	9826	9830	9834	9838	9842	9846	9850	9854	9857
2,2	9861	9864	9868	9871	9875	9878	9881	9884	9887	9890
2,3	9893	9896	9898	9901	9904	9906	9909	9911	9913	9916
2,4	9918	9920	9922	9925	9927	9929	9931	9932	9934	9936
2,5	9938	9940	9941	9943	9945	9946	9948	9949	9951	9952
2,6	9953	9955	9956	9957	9959	9960	9961	9962	9963	9964
2,7	9965	9966	9967	9968	9969	9970	9971	9972	9973	9974
2,8	9974	9975	9976	9977	9977	9978	9979	9979	9980	9981
2,9	9981	9982	9982	9983	9984	9984	9985	9985	9986	9986
3,0	9987	9987	9987	9988	9988	9989	9989	9989	9990	9990
3,1	9990	9991	9991	9991	9992	9992	9992	9992	9993	9993
3,2	9993	9993	9994	9994	9994	9994	9994	9995	9995	9995
3,3	9995	9995	9995	9996	9996	9996	9996	9996	9996	9997
3,4	9997	9997	9997	9997	9997	9997	9997	9997	9997	9998
3,5	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998
3,6	9998	9998	9999	9999	9999	9999	9999	9999	9999	9999
3,7	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
3,8	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999

Beispiele für den Gebrauch:

$\phi(2{,}32) = 0{,}9898$

$\phi(z) = 0{,}994 \Rightarrow z = 2{,}51$

$\phi(-0{,}9) = 1 - \phi(0{,}9) = 0{,}1841$

Tipps Stochastik Aufgabe HT 7

a) Klären Sie zunächst, welchem mathematischen Modell die Befragung der 800 Personen entspricht und legen Sie eine geeignete Zufallsgröße X fest. Die erwähnte Fehlentscheidung tritt in dem Fall ein, dass die Nullhypothese H_0 (im Aufgabentext «Hypothese» genannt) irrtümlicherweise abgelehnt wird (sog. Fehler 1. Art oder α-Fehler). Erläutern Sie, was das im vorliegenden Kontext bedeutet. Stellen Sie eine Ungleichung für den kritischen Wert $k \in \mathbb{N}$ auf, ab dem man die Nullhypothese verwirft. Rechnen Sie mit dem Gegenereignis; nutzen Sie als Näherung die Normalverteilung (überprüfen Sie vorher, dass die Bedingung $\sigma > 3$ erfüllt ist). Es gilt dabei $P(X \leqslant k) \approx \Phi\left(\frac{k+0,5-\mu}{\sigma}\right)$, wobei μ der Erwartungswert und σ die Standardabweichung von X ist und Φ die Gaußsche Integralfunktion, deren Werte in der Tabelle angegeben sind. Bestimmen Sie aus dem dort abgelesenen Wert den kritischen Wert k und den Ablehnungsbereich $\overline{A} = \{k+1, k+2, \ldots, 800\}$ von H_0.

b) Gehen Sie zur Beschreibung der möglichen Fehlentscheidung analog zu Aufgabenteil a) vor. Zur Berechnung der gesuchten Irrtumswahrscheinlichkeit berechnen Sie $P(X \leqslant 290)$ mit Hilfe der Näherung durch die Normalverteilung.
Gesucht ist weiterhin die Wahrscheinlichkeit des Fehlers, mit dem man irrtümlicherweise die Nullhypothese H_0: $p \geqslant 0,4$ annimmt, obwohl tatsächlich $p = 0,35$ gilt. Machen Sie sich klar, wann man H_0 Glauben schenkt und berechnen Sie die Wahrscheinlichkeit für diesen Fall unter der Bedingung, dass $p = 0,35$ ist. Nutzen Sie wiederum die Näherung durch die Normalverteilung und beachten Sie, dass sich μ und σ in Abhängigkeit von p ändern.

c) Erläutern Sie, warum es sich beim Ankreuzen bzw. Nicht-Ankreuzen der Kästchen einer bestimmten Zeile um eine Bernoulli-Kette handelt. Legen Sie eine geeignete Zufallsgröße Y fest und bestimmen Sie ihre Verteilung. Nutzen Sie zur Berechnung der gesuchten Wahrscheinlichkeiten in (1) bis (3) die Formel für binomialverteilte Zufallsgrößen Y: $P(Y = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$; $k \in \mathbb{N}$.

d) Betrachten Sie das Gegenereignis zu dem von Reinlich Senior erwähnten Ereignis. Nutzen Sie das Ergebnis aus Aufgabenteil c) (1) und berechnen Sie damit und mit dem Produktsatz die erste gesuchte Wahrscheinlichkeit. Klären Sie, wie die Aussagen von Reinlich Senior und Junior zueinander stehen, und berechnen Sie damit die zweite gesuchte Wahrscheinlichkeit.

e) Notieren Sie zunächst, welche bedingten und nicht-bedingten Wahrscheinlichkeiten gegeben sind und welche gesucht ist. Nutzen Sie entweder den Satz von Bayes oder zeichnen Sie ein Baumdiagramm, in dem die gegebenen Wahrscheinlichkeiten vorkommen und verwenden Sie die Definition der bedingten Wahrscheinlichkeit $P_B(A) = \frac{P(A \cap B)}{P(B)}$. Bestimmen Sie die darin vorkommenden Wahrscheinlichkeiten mit Hilfe des Baumdiagramms und der 1. und 2. Pfadregel.

Lösungen Stochastik Aufgabe HT 7

Vorüberlegungen: Bei der Befragung der 800 Personen sind nur die Antworten A: «Reinil ist mir bekannt» oder B: «Reinil ist mir nicht bekannt» möglich. Aufgrund der ersten Befragung kann davon ausgegangen werden, dass die Wahrscheinlichkeit, dass eine Person die Antwort A gibt, konstant ist. Damit handelt es sich bei der Befragung um eine Bernoulli-Kette. Ist X die Zufallsvariable, die die Anzahl der Antworten A bei 800 Befragten angibt, so ist X binomialverteilt mit $n = 800$ und $p = 0,4$ (gemäß der ersten Befragung).

a) Wenn Herr Reinlich Junior die Nullhypothese H_0: $p < 0,4$ überprüfen will, wird er diese nur dann verwerfen, wenn sehr viele der befragten Personen Antwort A geben. Dabei will er den Fehler dafür, dass er auf Grund der Befragung fälschlicherweise annimmt, Reinil ist bei 40 % oder mehr der Bevölkerung bekannt, gering halten (gemäß Aufgabenstellung soll diese Irrtumswahrscheinlichkeit für einen solchen Fehler 1. Art höchstens 5 % betragen).
Für die Entscheidungsregel bestimmt man den kritischen Wert $k \in \mathbb{N}; 0 \leqslant k \leqslant 800$ und damit den Ablehnungsbereich $\overline{A} = \{k+1, k+2, \ldots, 800\}$ von H_0 so, dass gilt:

$$P(X > k) \leqslant 0,05 \Leftrightarrow 1 - P(X \leqslant k) \leqslant 0,05 \Leftrightarrow P(X \leqslant k) \geqslant 0,95$$

Da für $n = 800$ keine Binomialverteilungstabelle existiert, muss man die Binomialverteilung durch die Normalverteilung annähern. Für die Standardabweichung σ von X gilt:

$$\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{800 \cdot 0,4 \cdot 0,6} = \sqrt{192} \approx 13,86 > 3$$

Deshalb ist die Näherung durch die Normalverteilung sinnvoll und es gilt:

$$P(X \leqslant k) \approx \Phi\left(\frac{k + 0,5 - \mu}{\sigma}\right)$$

wobei μ der Erwartungswert von X und Φ die Gaußsche Integralfunktion ist, deren Werte in der Tabelle angegeben sind. Da für den Erwartungswert von X gilt: $\mu = n \cdot p = 320$, folgt:

$$P(X \leqslant k) \approx \Phi\left(\frac{k + 0,5 - 320}{\sqrt{192}}\right) \geqslant 0,95$$

Mit Hilfe der Normalverteilungstabelle findet man $\Phi(1,64) = 0,9495$ und $\Phi(1,65) = 0,9505$. Also gilt für das Argument von Φ: $\frac{k+0,5-320}{\sqrt{192}} = 1,65$. Umformen der Gleichung nach k führt zu $k \approx 342,36$. Also liegt der kritische Wert bei $k = 343$ und für den Ablehnungsbereich von H_0 gilt: $\overline{A} = \{344, 345, \ldots, 800\}$.
Geben von den befragten Personen also 344 oder mehr die Antwort A, d.h. sie sagen, dass sie Reinil kennen, wird man der Gegenhypothese, dass der Bekanntheitsgrad des Waschmittels 40% oder mehr beträgt, Glauben schenken.

b) Wenn Herr Reinlich Senior seine Nullhypothese H_0: $p \geqslant 0,4$ überprüfen will, wird er diese nur dann verwerfen, wenn sehr wenige befragte Personen die Antwort A geben. Dabei will er den Fehler dafür, dass er auf Grund der Befragung fälschlicherweise annimmt, Reinil ist bei weniger als 40 % der Bevölkerung bekannt, gering halten. Die zugehörige Irrtumswahrscheinlichkeit α, die zu der Entscheidungsregel «Ablehnung von $H_0 : p \geq 0,4$, wenn höchstens 290 der Befragten Reinil kennen» gehört, soll im folgenden errechnet werden:

$$\alpha = P(X \leq 290) \approx \Phi\left(\frac{290+0,5-320}{\sqrt{192}}\right) \approx \Phi(-2,13) = 1-\Phi(2,13) = 0,0166 = 1,66\%$$

Also wird man in ca. 1,66 % der Fälle fälschlicherweise annehmen, dass der Bekanntheitsgrad von Reinil unter 40 % liegt.

Falls der Bekanntheitsgrad nun tatsächlich nur 35 % beträgt, macht man einen Fehler, wenn man weiterhin der Nullhypothese H_0: $p \geq 0,4$ glaubt. Dies ist genau dann der Fall, wenn mehr als 290 der Befragten Antwort A geben. Für diese Wahrscheinlichkeit eines sogenannten Fehlers 2. Art berechnet man $P(X > 290)$. Dabei ist X nun aber binomialverteilt mit $n = 800$ und $p = 0,35$. Daraus folgt, dass nun der Erwartungswert $\mu = 800 \cdot 0,35 = 280$ und die Standardabweichung $\sigma = \sqrt{800 \cdot 0,35 \cdot 0,65} = \sqrt{182} \approx 13,49$ ist. Insbesondere ist $\sigma > 3$ und wie oben kann die Binomialverteilung durch die Normalverteilung approximiert werden:

$$P(X > 290) = 1 - P(X \leq 290) \approx 1 - \Phi\left(\frac{290+0,5-280}{\sqrt{182}}\right) \approx 1 - \Phi(0,78) = 1 - 0,7823$$
$$= 0,2177 = 21,77\%$$

Der Wert von Φ ist dabei der Tabelle entnommen.

Die Entscheidungsregel führt also bei einem tatsächlichen Bekanntheitsgrad von 35 % in ca. 21,8 % der Fälle zu einer falschen Entscheidung.

c) In jeder der 100 Zeilen befinden sich 8 Kästchen. Die Wahrscheinlichkeit, dass ein Kästchen angekreuzt wird, beträgt $p = 0,4$; die Wahrscheinlichkeit, dass es nicht angekreuzt wird, ist demzufolge $1 - p = 0,6$. Da nur diese beiden Fälle eintreten können, handelt es sich in jeder Zeile beim Ausfüllen der 8 Kästchen um eine Bernoulli-Kette. Sei Y die Zufallsvariable, die die Anzahl angekreuzter Kästchen in einer Zeile des Bogens beschreibt. Dann ist Y binomialverteilt mit $n = 8$ und $p = 0,4$, d.h. man kann die Formel für binomialverteilte Zufallsvariablen benutzen:

(1) $P(Y = 0) = \binom{8}{0} \cdot 0,4^0 \cdot 0,6^8 = 0,6^8 \approx 0,0168 = 1,68\%$

(2) $P(Y = 4) = \binom{8}{4} \cdot 0,4^4 \cdot 0,6^4 \approx 0,2322 = 23,22\%$

(3) Hier kann man mit dem Gegenereignis rechnen, um eine Einzelwahrscheinlichkeit weniger bestimmen zu müssen:

$$P(Y \geq 4) = 1 - P(Y = 0) - P(Y = 1) - P(Y = 2) - P(Y = 3)$$
$$= 1 - 0,6^8 - \binom{8}{1} \cdot 0,4^1 \cdot 0,6^7 - \binom{8}{2} \cdot 0,4^2 \cdot 0,6^6 - \binom{8}{3} \cdot 0,4^3 \cdot 0,6^5$$
$$\approx 0,4059 = 40,59\%$$