Binomialverteilung, Gaußfunktion, Näherungsformel von de Moivre - LaPlace, Test- und Schätzverfahren

Arno Fehringer, April 2017

Quellen

Büchter, A.: Henn, H. : Elementare Stochastik, Springer Verlag Berlin Heidelberg 2005

Athen, H. : Wahrscheinlichkeitsrechnung und Statistik, Verlag Schroedel Schöningh 1968, 2. Aufl.

Lambacher Schweizer : Stochastik
genauer:
Schmid, A. ; Stark, J. : Stochastik Leistungskurs, Klett Verlag Stuttgart 1995, 1. Aufl.

Binomialverteilungen $B_{n\,;\,p}(k)$

Ein Zufallsexperiment mit genau zwei Ausgängen „Treffer“ und „Fehlschlag“ und den jeweiligen Wahrscheinlichkeiten p und q = 1– p werde n-mal wiederholt.

Die Wahrscheinlichkeit für k „Treffer“ (und demzufolge n-k „Fehlschlägen“) , k = 0, 1, 2, … , n , ist gegeben durch

$$p(k) \;=\; \binom{n}{k}\, p^k (1-p)^{n-k}$$

Die Wahrscheinlichkeitsverteilung heißt **Binomialverteilung $B_{n\,;\,p}(k)$** .

Man schreibt auch manchmal $B_{n;p}(k) \;:=\; p(k) \;=\; \binom{n}{k}\, p^k (1-p)^{n-k}$.

Der **Erwartungswert** ist :

$$E \;=\; \sum_{k=0}^{n} k \cdot p(k)$$

$$E \;=\; \sum_{k=0}^{n} k \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

$$E \;=\; \sum_{k=1}^{n} k \cdot \frac{n!}{k!\cdot(n-k)!} \cdot p^k \cdot (1-p)^{n-k}$$

$$E \;=\; np \cdot \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\cdot(n-k)!} \cdot p^{k-1} \cdot (1-p)^{n-k}$$

$$E \;=\; np \cdot \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\cdot(n-1-(k-1))!} \cdot p^{k-1} \cdot (1-p)^{n-1-(k-1)}$$

$$l \;:=\; k-1 \quad , \quad l \;=\; 0 \;,\; \ldots \;,\; n-1$$

$$E \;=\; np \cdot \sum_{l=0}^{n-1} \frac{(n-1)!}{l!\cdot(n-1-l)!} \cdot p^{l} \cdot (1-p)^{n-1-l}$$

$$E \;=\; np \cdot \sum_{l=0}^{n-1} \binom{n-1}{l} \cdot p^{l} \cdot (1-p)^{n-1-l}$$

$$E \;=\; np \cdot (p \;+\; 1-p)^{n-1}$$

$$E \;=\; np \cdot 1^{n-1}$$

$$\boxed{E \;=\; np} \quad .$$

Die **Varianz** ist :

$$V = \sum_{k=0}^{n} (k - E)^2 \cdot p(k)$$

$$V = \sum_{k=0}^{n} (k^2 - 2kE + E^2) \cdot p(k)$$

$$V = \sum_{k=0}^{n} k^2 \cdot p(k) - 2E \cdot \sum_{k=0}^{n} k \cdot p(k) + \sum_{k=0}^{n} E^2 \cdot p(k)$$

$$V = \sum_{k=0}^{n} k^2 \cdot p(k) - 2E \cdot E + E^2 \cdot \sum_{k=0}^{n} p(k)$$

$$V = \sum_{k=0}^{n} k^2 \cdot p(k) - 2E^2 + E^2 \cdot 1$$

$$V = \sum_{k=1}^{n} k^2 \cdot p(k) - E^2$$

$$V = \sum_{k=1}^{n} k^2 \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \qquad - \qquad n^2 p^2$$

$$V = \sum_{k=1}^{n} k^2 \cdot \frac{n!}{k! \cdot (n-k)!} \cdot p^k \cdot (1-p)^{n-k} \qquad - \qquad n^2 p^2$$

$$V = np \cdot \sum_{k=1}^{n} k \cdot \frac{(n-1)!}{(k-1)! \cdot (n-k)!} \cdot p^k \cdot (1-p)^{n-k} \qquad - \qquad n^2 p^2$$

$$V = np \cdot \sum_{k=1}^{n} k \cdot \frac{(n-1)!}{(k-1)! \cdot (n-1-(k-1))!} \cdot p^{k-1} \cdot (1-p)^{n-1-(k-1)} \qquad - \qquad n^2 p^2$$

$$l := k-1 \quad , \quad l = 0 \, , \, \ldots \, , \, n-1$$

$$V = np \cdot \sum_{l=0}^{n-1} (l+1) \cdot \frac{(n-1)!}{l! \cdot (n-1-l)!} \cdot p^l \cdot (1-p)^{n-1-l} \qquad - \qquad n^2 p^2$$

$$V = np \cdot \left(\sum_{l=0}^{n-1} l \cdot \frac{(n-1)!}{l! \cdot (n-1-l)!} \cdot p^l \cdot (1-p)^{n-1-l} + \sum_{l=0}^{n-1} \frac{(n-1)!}{l! \cdot (n-1-l)!} \cdot p^l \cdot (1-p)^{n-1-l} \right) - n^2 p^2$$

$$V = np \cdot \left(\sum_{l=0}^{n-1} l \cdot \binom{n-1}{l} \cdot p^l \cdot (1-p)^{n-1-l} + \sum_{l=0}^{n-1} \binom{n-1}{l} \cdot p^l \cdot (1-p)^{n-1-l} \right) - n^2 p^2$$

$$V = np \cdot \left((n-1)p + (p + 1-p)^{n-1} \right) - n^2p^2$$

$$V = np \cdot (np - p + 1) - n^2p^2$$

$$V = np \cdot (np - p + 1 - np)$$

$$\boxed{V = np(1-p)} \, .$$

Die **Standardabweichung** ist :

$$\sigma = \sqrt{V}$$

$$\boxed{\sigma = \sqrt{np(1-p)}} \, .$$

Zusammenfassung für die **Binomialverteilung $B_{n\,;\,p}(k)$** :

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad , \quad k = 0 \, , \, \ldots \, , \, n$$

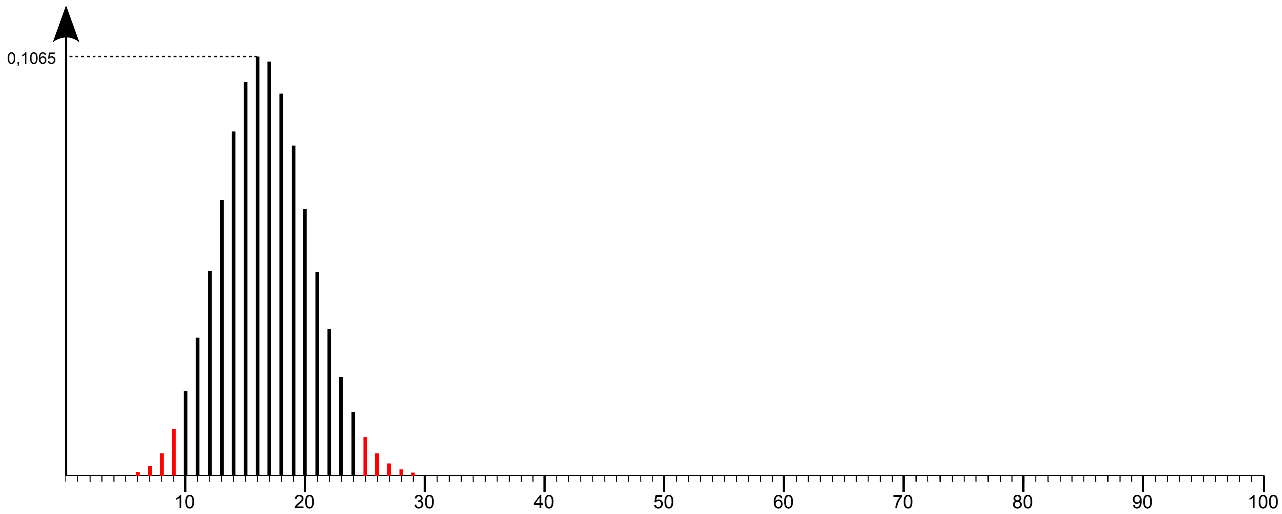
$$E = np$$

$$V = np(1-p)$$

$$\sigma = \sqrt{np(1-p)}$$

Binomialverteilung für das Würfelexperiment „ k-mal die 6 “ bei 100 Würfen

$B_{100\,;\,1/6}(k)$



Intervallwahrscheinlichkeiten

$$P\left(k_1 \le k \le k_2\right) = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k \left(1-p\right)^{n-k}$$

Zur Berechnung nimmt man die **Summenfunktion der Binomialverteilung $\mathbf{F_{n\,;\,p}(k)}$** .
Sie ist definiert als

$$F_{n;p}(k) := \sum_{k=0_1}^{k} \binom{n}{k} p^k \left(1-p\right)^{n-k} .$$

Damit ergibt sich :

$$P\left(k_1 \le k \le k_2\right) = \sum_{k=0}^{k_2} \binom{n}{k} p^k \left(1-p\right)^{n-k} - \sum_{k=0}^{k_1} \binom{n}{k} p^k \left(1-p\right)^{n-k}$$

$$P\left(k_1 \le k \le k_2\right) = F_{n;p}\left(k_2\right) - F_{n;p}\left(k_1\right)$$

Die **Summenfunktionen der Binomialverteilungen $\mathbf{F_{n\,;\,p}(k)}$** liegen tabelliert vor (siehe nächste Seite) .

Tabelle V: Binomialverteilung

Summenfunktion

$F_{n;p}(k) = B_{n;p}(0) + \ldots + B_{n;p}(k)$

n	k	p: 0,02	0,03	0,04	0,05	0,10	1/6	0,20	0,30	1/3	0,40	0,50	k	n
	0	0,1326	0476	0169	0059	0000	0000	0000	0000	0000	0000	0000	99	
	1	4033	1946	0872	0371	0003	0000	0000	0000	0000	0000	0000	98	
	2	6767	4198	2321	1183	0019	0000	0000	0000	0000	0000	0000	97	
	3	8590	6472	4295	2578	0078	0000	0000	0000	0000	0000	0000	96	
	4	9492	8179	6289	4360	0237	0001	0000	0000	0000	0000	0000	95	
	5	9845	9192	7884	6160	0576	0004	0000	0000	0000	0000	0000	94	
	6	9959	9688	8936	7660	1172	0013	0001	0000	0000	0000	0000	93	
	7	9991	9894	9525	8720	2061	0038	0003	0000	0000	0000	0000	92	
	8	9998	9968	9810	9369	3209	0095	0009	0000	0000	0000	0000	91	
	9		9991	9932	9718	4513	0213	0023	0000	0000	0000	0000	90	
	10		9998	9978	9885	5832	0427	0057	0000	0000	0000	0000	89	
	11			9993	9957	7030	0777	0126	0000	0000	0000	0000	88	
	12			9998	9985	8018	1297	0253	0000	0000	0000	0000	87	
	13				9995	8761	2000	0469	0001	0000	0000	0000	86	
	14				9999	9274	2874	0804	0002	0000	0000	0000	85	
	15					9601	3877	1285	0004	0000	0000	0000	84	
	16					9794	4942	1923	0010	0001	0000	0000	83	
	17					9900	5994	2712	0022	0002	0000	0000	82	
	18					9954	6965	3621	0045	0005	0000	0000	81	
	19					9980	7803	4602	0089	0011	0000	0000	80	
	20					9992	8481	5595	0165	0024	0000	0000	79	
	21					9997	8998	6540	0288	0048	0000	0000	78	
	22					9999	9370	7389	0479	0091	0001	0000	77	
	23						9621	8109	0755	0164	0003	0000	76	
	24						9783	8686	1136	0281	0006	0000	75	
	25						9881	9125	1631	0458	0012	0000	74	
	26						9938	9442	2244	0715	0024	0000	73	
	27						9969	9658	2964	1066	0046	0000	72	
	28						9985	9800	3768	1524	0084	0000	71	
	29						9993	9888	4623	2093	0148	0000	70	
	30						9997	9939	5491	2766	0248	0000	69	
	31						9999	9969	6331	3525	0398	0001	68	
	32							9985	7107	4344	0615	0002	67	
	33							9993	7793	5188	0913	0004	66	
	34							9997	8371	6019	1303	0009	65	
	35							9999	8839	6803	1795	0018	64	
100	36							9999	9201	7511	2386	0033	63	100
	37								9470	8123	3068	0060	62	
	38								9660	8630	3822	0105	61	
	39								9790	9034	4621	0176	60	
	40								9875	9341	5433	0284	59	
	41								9928	9566	6225	0443	58	
	42								9960	9724	6967	0666	57	
	43								9979	9831	7635	0967	56	
	44								9989	9900	8211	1356	55	
	45								9995	9943	8689	1841	54	
	46								9997	9969	9070	2421	53	
	47								9999	9983	9362	3087	52	
	48								9999	9991	9577	3822	51	
	49									9996	9729	4602	50	
	50									9998	9832	5398	49	
	51									9999	9900	6178	48	
	52										9942	6914	47	
	53										9968	7579	46	
	54										9983	8159	45	
	55										9991	8644	44	
	56										9996	9033	43	
	57										9998	9334	42	
	58										9999	9557	41	
	59											9716	40	
	60											9824	39	
	61											9895	38	
	62											9940	37	
	63											9967	36	
	64											9982	35	
	65											9991	34	
	66											9996	33	
	67											9998	32	
	68											9999	31	
n		p: 0,98	0,97	0,96	0,95	0,90	5/6	0,80	0,70	2/3	0,60	0,50	k	n

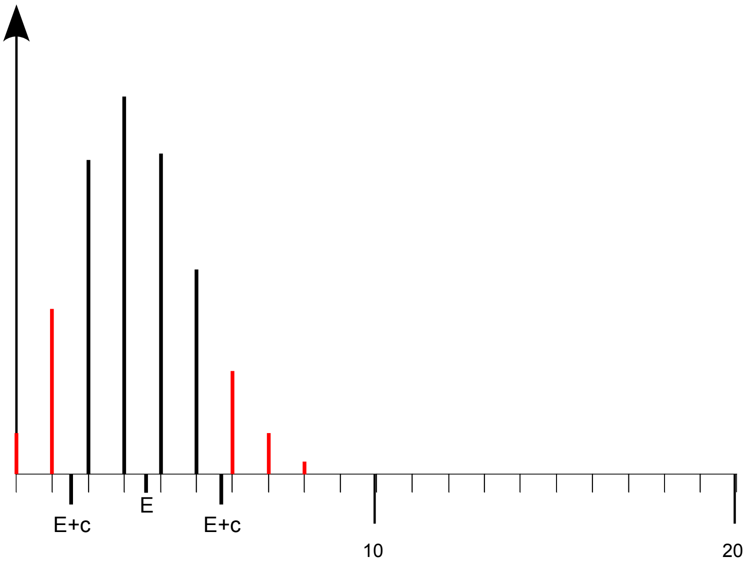
Bei rot gedrucktem Eingang, d. h. $p \geqq 0{,}5$ gilt:
$F_{n;p}(k) = 1 -$ abgelesener Wert

Nicht aufgeführte Werte sind (auf 4 Dez.) 1,0000.

Tschebyscheff – Ungleichung

Pafnuti Lwowitsch Tschebyscheff , russischer Mathematiker (1821 – 1894)

$B_{20\,;\,1/6}(k) \qquad E = 20 \cdot \frac{1}{6} = 3{,}\overline{6}$



$$p(|k-E| \geq c) = \sum_{|k-E|\geq c}^{n} p(k) = ?$$

$$\sigma^2 = V = \sum_{k=0}^{n} (k-E)^2 \cdot p(k)$$

$$\sigma^2 = \sum_{|k-E|<c}^{n} (k-E)^2 \cdot p(k) + \sum_{|k-E|\geq c}^{n} (k-E)^2 \cdot p(k)$$

$$\sigma^2 \geq \sum_{|k-E|\geq c}^{n} (k-E)^2 \cdot p(k)$$

$$\sigma^2 \geq \sum_{|k-E|\geq c}^{n} c^2 \cdot p(k)$$

$$\sigma^2 \geq c^2 \cdot \sum_{|k-E|\geq c}^{n} p(k)$$

$$\boxed{p(|k-E| \geq c) = \sum_{|k-E|\geq c}^{n} p(k) \leq \frac{\sigma^2}{c^2}}$$

Tschebyscheff – Ungleichung

Wählt man für c Vielfache von σ, also $c = 1\sigma$, 2σ , 3σ , erhält man

$$p(|k-E| \geq 1\sigma) \leq 1$$

$$p(|k-E| \geq 2\sigma) \leq \frac{1}{4}$$

$$\underline{p(|k-E| \geq 2\sigma) \leq 25\%}$$

$$p(|k-E| \geq 3\sigma) \leq \frac{1}{9}$$

$$\underline{p(|k-E| \geq 3\sigma) \leq 11.2\%}$$

$$\Rightarrow \quad -p(|k-E| \geq 2\sigma) \geq -\frac{1}{4}$$

$$1 - p(|k-E| \geq 2\sigma) \geq 1 - \frac{1}{4}$$

$$1 - p(|k-E| \geq 2\sigma) \geq \frac{3}{4}$$

$$\underline{1 - p(|k-E| < 2\sigma) \geq 75\%}$$

$$\Rightarrow \quad -p(|k-E| \geq 3\sigma) \geq -\frac{1}{9}$$

$$1 - p(|k-E| \geq 3\sigma) \geq 1 - \frac{1}{9}$$

$$1 - p(|k-E| \geq 3\sigma) \geq \frac{8}{9}$$

$$\underline{1 - p(|k-E| < 3\sigma) \geq 88.8\%}$$

Relative Häufigkeit für „Treffer“ und das Gesetz der großen Zahlen

Bei einer $B_{n\,;\,p}$ - verteilten Zufallsvariablen mit $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$ gibt die Variable $\bar{k} = \frac{k}{n}$ die **relative Häufigkeit für „Treffer“** an . Es ist $p(\bar{k}) = p(k)$.

Der **Erwartungswert** ist

$$E(\bar{k}) = \sum_{k=0}^{n} \bar{k} \cdot p(\bar{k})$$

$$E(\bar{k}) = \sum_{k=0}^{n} \frac{k}{n} \cdot p(k)$$

$$E(\bar{k}) = \sum_{k=0}^{n} \frac{k}{n} \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

$$E(\bar{k}) = \frac{1}{n} \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

$$E(\bar{k}) = \frac{1}{n} E$$

$$E(\bar{k}) = \frac{1}{n} np$$

$$\boxed{E(\bar{k}) = p} \quad .$$

Die **Varianz** ist

$$V(\bar{k}) = \sum_{k=0}^{n} \left(\bar{k} - E(\bar{k})\right)^2 \cdot p(\bar{k})$$

$$V(\bar{k}) = \sum_{k=0}^{n} \left(\bar{k} - p\right)^2 \cdot p(k)$$

$$V(\bar{k}) = \sum_{k=0}^{n} \left(\overline{k^2} - 2\bar{k}p + p^2\right) \cdot p(k)$$

$$V(\bar{k}) = \sum_{k=0}^{n} \overline{k^2} \cdot p(k) - \sum_{k=0}^{n} 2\bar{k}p \cdot p(k) + \sum_{k=0}^{n} p^2 \cdot p(k)$$

$$V(\bar{k}) = \sum_{k=1}^{n} \overline{k^2} \cdot p(k) - \sum_{k=0}^{n} 2\bar{k}p \cdot p(k) + \sum_{k=0}^{n} p^2 \cdot p(k)$$

$$V(\bar{k}) = \sum_{k=1}^{n} \frac{k^2}{n^2} \cdot p(k) - 2p \sum_{k=0}^{n} \bar{k} \cdot p(\bar{k}) + p^2 \sum_{k=0}^{n} p(k)$$

$$V(\bar{k}) = \sum_{k=1}^{n} \frac{k^2}{n^2} \cdot p(k) \quad - 2p\ E(\bar{k}) + p^2 \cdot 1$$

$$V(\bar{k}) = \sum_{k=1}^{n} \frac{k^2}{n^2} \cdot p(k) \quad - 2p^2 + p^2$$

$$V(\bar{k}) = \sum_{k=1}^{n} \frac{k^2}{n^2} \cdot p(k) \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2} \sum_{k=1}^{n} k^2 \cdot p(k) \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2} \sum_{k=1}^{n} k^2 \cdot \frac{n!}{k!\,(n-k)!}\ p^k\ (1-p)^{n-k} \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\ np \sum_{k=1}^{n} k \cdot \frac{(n-1)!}{(k-1)!\,(n-k)!}\ p^{k-1}\ (1-p)^{n-k} \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\ np \sum_{k=1}^{n} k \cdot \frac{(n-1)!}{(k-1)!\,(n-1-(k-1))!}\ p^{k-1}\ (1-p)^{n-1-(k-1)} \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\ np \sum_{k=1}^{n} k \cdot \frac{(n-1)!}{(k-1)!\,(n-1-(k-1))!}\ p^{k-1}\ (1-p)^{n-1-(k-1)} \quad - p^2$$

$$l := k-1 \quad , \quad l = 0\ , \ldots ,\ n-1$$

$$V(\bar{k}) = \frac{1}{n^2}\ np \sum_{l=0}^{n-1} (l+1) \cdot \frac{(n-1)!}{l!\,(n-1-l)!}\ p^l\ (1-p)^{n-1-l} \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\ np \left(\sum_{l=0}^{n-1} l \cdot \frac{(n-1)!}{l!\,(n-1-l)!}\ p^l\ (1-p)^{n-1-l} + 1 \cdot \frac{(n-1)!}{l!\,(n-1-l)!}\ p^l\ (1-p)^{n-1-l} \right) \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\ np \left(\sum_{l=0}^{n-1} l \cdot \frac{(n-1)!}{l!\,(n-1-l)!}\ p^l\ (1-p)^{n-1-l} + \sum_{l=0}^{n-1} \frac{(n-1)!}{l!\,(n-1-l)!}\ p^l\ (1-p)^{n-1-l} \right) \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\ np \left(\ (n-1)p + 1 \ \right) \quad - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\, np\left(np - p + 1 \right) - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\left(n^2p^2 - np^2 + np \right) - p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\left(n^2p^2 - np^2 + np \right) - \frac{1}{n^2}\, n^2p^2$$

$$V(\bar{k}) = \frac{1}{n^2}\left(n^2p^2 - np^2 + np - n^2p^2 \right)$$

$$V(\bar{k}) = \frac{1}{n^2}\left(- np^2 + np \right)$$

$$V(\bar{k}) = \frac{1}{n^2}\left(np - np^2 \right)$$

$$V(\bar{k}) = \frac{1}{n^2}\, np(1 - p)$$

$$V(\bar{k}) = \frac{1}{n^2}\, V$$

$$\boxed{V(\bar{k}) = \frac{1}{n}\, p(1 - p)} \quad .$$

Die **Standardabweichung** ist

$$\boxed{\sigma(\bar{k}) = \sqrt{\frac{1}{n}\, p(1 - p)}}$$

Die Tschebischeff – Ungleichung für die Variable $\bar{k} = \frac{k}{n}$ liefert

$$p\left(\left|\, \bar{k} - E(\bar{k}) \,\right| \geq c \right) \leq \frac{\sigma(\bar{k})^2}{c^2}$$

$$p\left(\left|\, \bar{k} - p \,\right| \geq c \right) \leq \frac{\frac{1}{n}\, p(1 - p)}{c^2}$$

$$p(|\,\bar{k} - p\,| \geq c\,) \;\leq\; \frac{p(1 - p)}{nc^2}$$

$$p(|\,\bar{k} - p\,| < c\,) \;=\; 1 - p(|\,\bar{k} - p\,| \geq c\,)$$

$$p(|\,\bar{k} - p\,| < c\,) \;\geq\; 1 - \frac{p(1 - p)}{nc^2}$$

Wegen

$$-p(1-p) \;=\; -p + p^2$$

$$-p(1-p) \;=\; p^2 - p$$

$$-p(1-p) \;=\; p^2 - 2\cdot p\cdot\frac{1}{2}$$

$$-p(1-p) \;=\; p^2 - 2\cdot p\cdot\frac{1}{2} + \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2$$

$$-p(1-p) \;=\; \left(p - \frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2$$

$$-p(1-p) \;\geq\; -\left(\frac{1}{2}\right)^2$$

$$-p(1-p) \;\geq\; -\frac{1}{4}$$

folgt

$$p(|\,\bar{k} - p\,| < c\,) \;\geq\; 1 - \frac{p(1 - p)}{nc^2}$$

$$p(|\,\bar{k} - p\,| < c\,) \;\geq\; 1 - \frac{1}{4nc^2}$$

Satz vom Gesetz der großen Zahlen

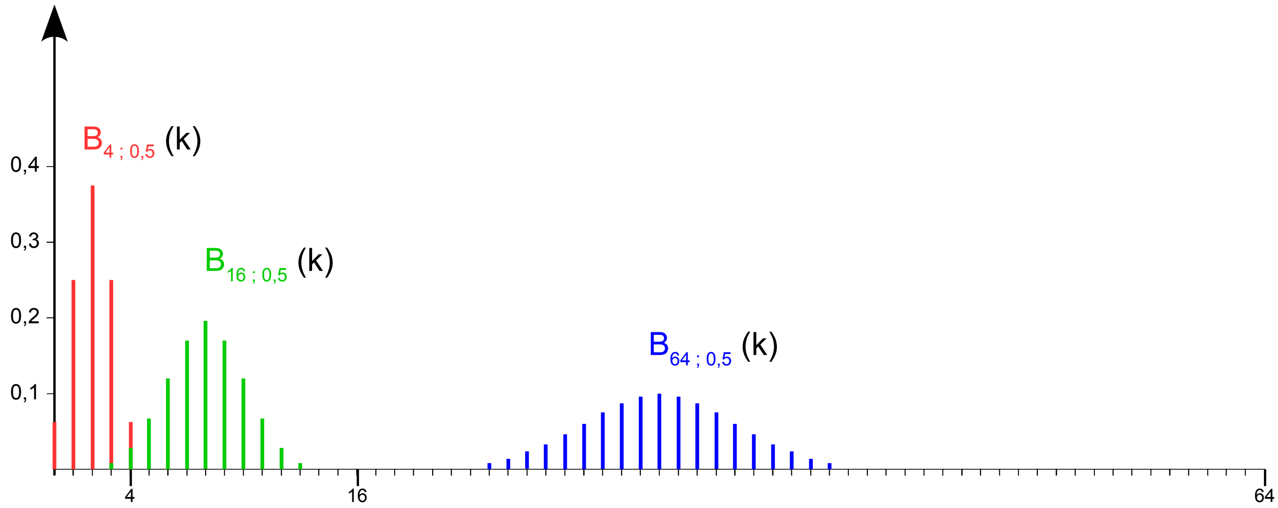
Jakob Bernoulli, schweizer Mathematiker (1654 – 1705)

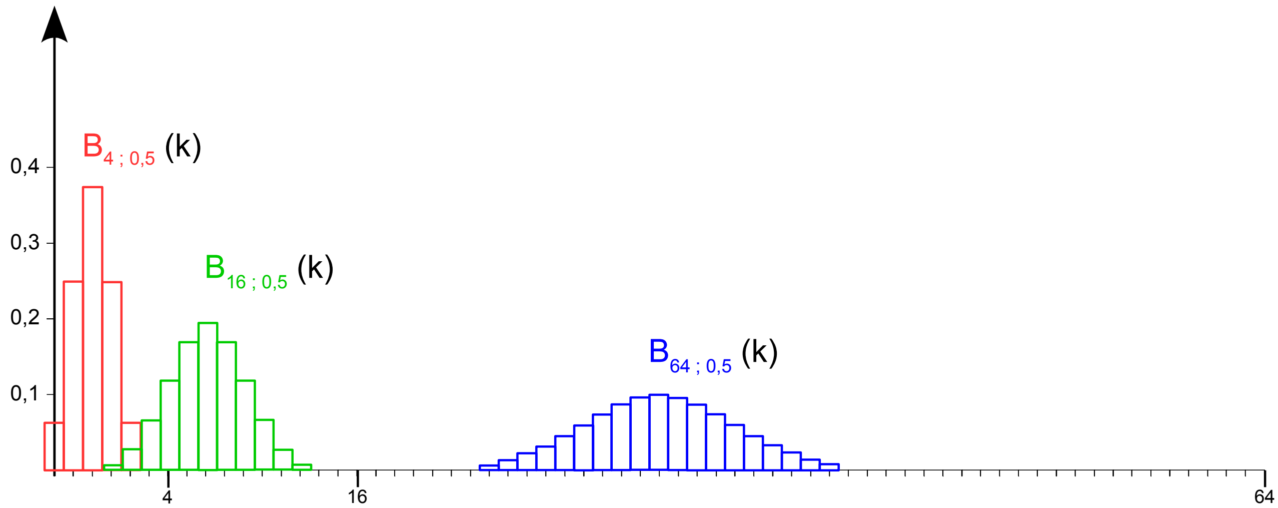
Die Wahrscheinlichkeit, dass die relative Häufigkeit für „Treffer“ um weniger als c von der Wahrscheinlichkeit p abweicht, strebt für $n \longrightarrow \infty$ gegen 1.

$$p(|\,\bar{k} - p\,| < c\,) \;\geq\; 1 - \frac{1}{4nc^2}$$

Verhalten der Binomialverteilung $B_{n\,;\,p}(k)$ für große n

Schaubilder bzw. Histogramme der Binomialverteilungen
$B_{4\,;\,0,5}(k)$, $B_{16\,;\,0,5}(k)$, $B_{64\,;\,0,5}(k)$, … , $B_{2^{2l}\,;\,0,5}(k)$, … mit l = 1, 2, 3, ...





Die Folge der Standardabweichungen $\sigma = \sqrt{np(1-p)} = 0.5\sqrt{n}$ verdoppeln sich :

$\sigma_4 = 1$, $\sigma_{16} = 2$, $\sigma_{64} = 4$, ...

Die Folge der Maxima halbieren sich in etwa :

$M_4 = 0,375$, $M_{16} = 0,196$, $M_{64} = 0,099$, ...

Die Folge der Erwartungswerte $E = np$ driftet nach rechts :

$E_4 = 2$, $E_{16} = 8$, $E_{64} = 32$, ...

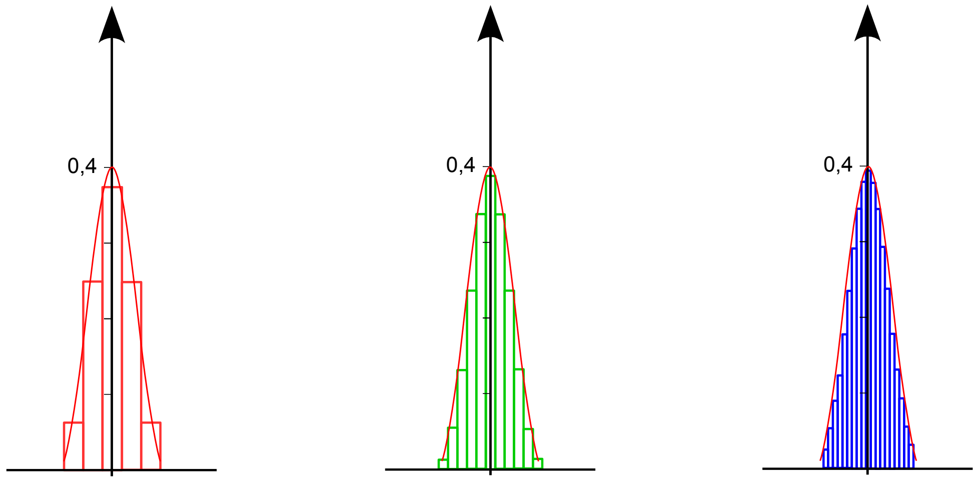
Betrachtet man die Histogramme der Binomialverteilungen $B_{n\,;\,p}(k)$ für wachsendes n und konstantes p , so werden diese immer breiter aber flacher, und die Erwartungswerte wandern nach rechts .
Die Wahrscheinlichkeit eines einzelnen Ereignisses

$$p(k) \;=\; \binom{n}{k}\, p^{k}(1-p)^{n-k}$$ wird immer kleiner ,

da die Flächensumme $\sum_{k=0}^{n} \binom{n}{k}\, p^{k}(1-p)^{n-k} \;=\; 1$ ist .

Um das Verhalten der Binomialverteilungen $B_{n\,;\,p}(k)$ für große n besser untersuchen zu können, verschiebt man die Histogramme um den Erwartungswert $E = np$ nach links, staucht die Werte von k und die Rechteckbreiten mit dem Faktor $\frac{1}{\sigma}$ und streckt die Funktionswerte p und die Rechteckhöhen mit dem Faktor σ , wobei $\sigma = \sqrt{np(1-p)}$ ist . Diesen Prozess nennt man **Normierung** .

Die Flächensummen bleiben weiterhin gleich 1 .



Dadurch erhält man eine neue Zufallsvariable $x_k = \frac{k-E}{\sigma}$

mit der Wahrscheinlichkeit

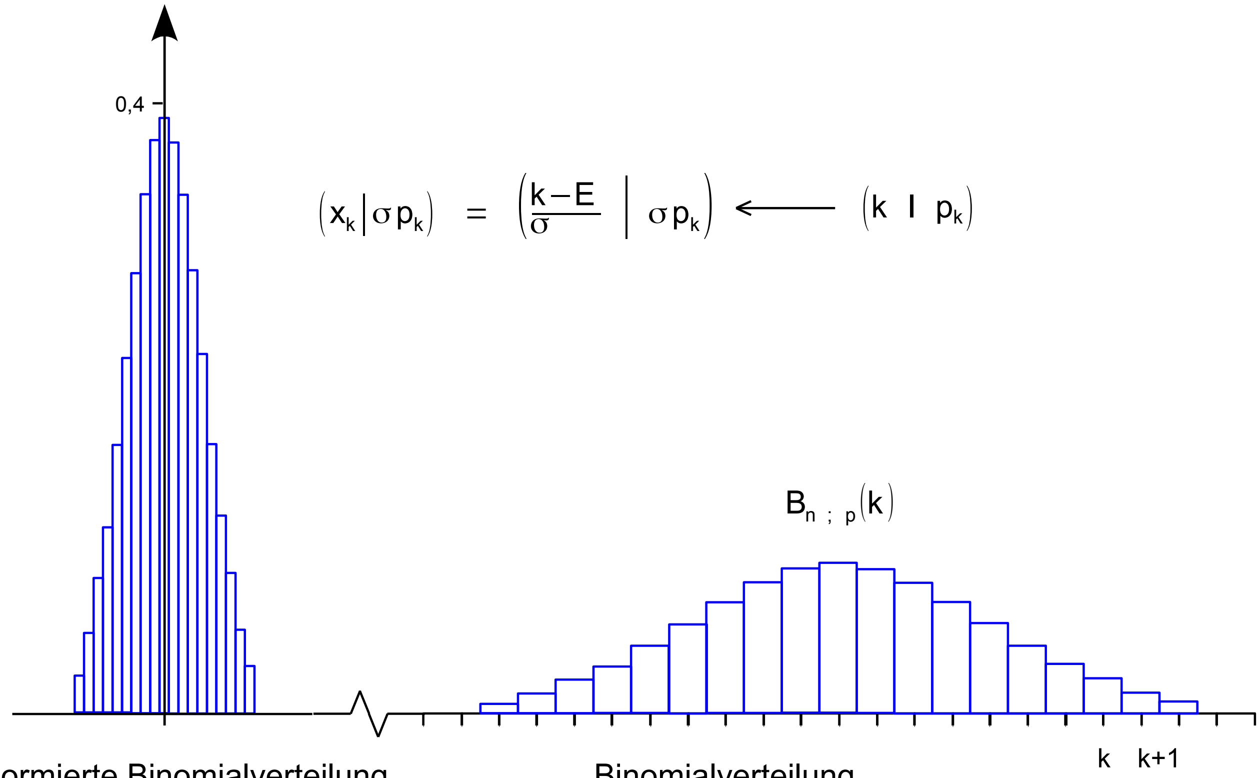
$$p(x_k) \;=\; p\left(\frac{k-E}{\sigma}\right) \;=\; \sigma \cdot p(k) \;=\; \sigma \cdot \binom{n}{k}\, p^{k}(1-p)^{n-k} \quad ,$$

wobei $E = np$ und $\sigma = \sqrt{np(1-p)}$ ist,

und entsprechende Histogramme.

Die Folge Histogramme der so normierten Binomialverteilungen scheinen sich immer besser an eine glockenförmige **Grenzfunktion** $\phi(x)$ anzuschmiegen .

Die **Gaußsche Glockenfunktion** als Grenzfunktion der normierten Binomialverteilungen



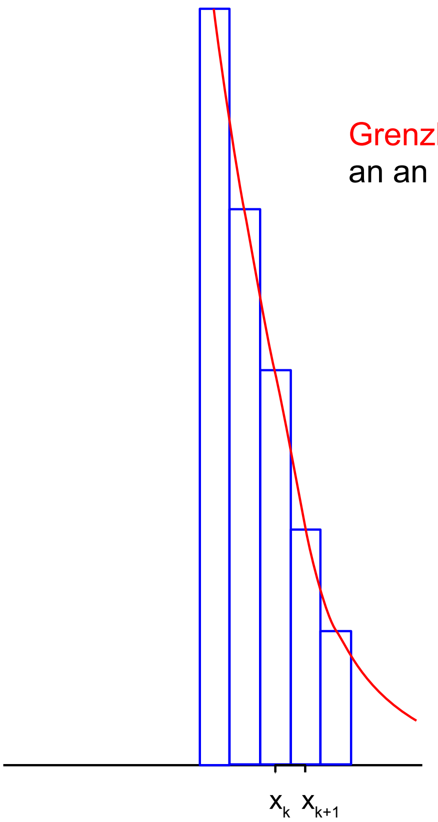
$$\left(x_k \,\middle|\, \sigma p_k\right) = \left(\frac{k-E}{\sigma} \,\middle|\, \sigma p_k\right) \longleftarrow \left(k \,\middle|\, p_k\right)$$

Normierte Binomialverteilung

$$g_n(x_k) = g_n\left(\frac{k-E}{\sigma}\right) = \sigma p_k$$

Binomialverteilung

$$p_k = p(k) = \binom{n}{k} p^k q^{n-k} \quad , \quad q = 1-p$$



Grenzkurve ϕ
an an die sich die Funktionen g_n annähern

Relative Änderung der Funktion g_n :

$$\frac{\Delta g_n}{g_n} = \frac{g_n(x_{k+1}) - g_n(x_k)}{g_n(x_k)}$$

$$\frac{\Delta g_n}{g_n} = \frac{\sigma p_{k+1} - \sigma p_k}{\sigma p_k}$$

$$\frac{\Delta g_n}{g_n} = \frac{p_{k+1} - p_k}{p_k}$$

Wegen

$$\binom{n}{k+1} = \frac{n!}{(k+1)!(n-k-1)!}$$

$$\binom{n}{k+1} = \frac{n!(n-k)}{(k+1)k!(n-k)!}$$

$$\binom{n}{k+1} = \frac{(n-k)}{(k+1)}\,\frac{n!}{k!(n-k)!}$$

ist

$$p_{k+1} = \binom{n}{k+1} p^{k+1} q^{n-k-1}$$

$$p_{k+1} = \frac{(n-k)}{(k+1)}\,\frac{n!}{k!(n-k)!}\,p^{k+1} q^{n-k-1}$$

$$p_{k+1} = \frac{(n-k)}{(k+1)}\,\frac{p}{q}\,\frac{n!}{k!(n-k)!}\,p^{k} q^{n-k}$$

$$p_{k+1} = \frac{(n-k)}{(k+1)}\,\frac{p}{q}\,p_k$$

und es folgt

$$\frac{\Delta g_n}{g_n} = \frac{p_{k+1} - p_k}{p_k}$$

$$\frac{\Delta g_n}{g_n} = \frac{\frac{(n-k)}{(k+1)}\,\frac{p}{q}\,p_k - p_k}{p_k}$$

$$\frac{\Delta g_n}{g_n} = \frac{(n-k)}{(k+1)}\frac{p}{q} - 1$$

$$\frac{\Delta g_n}{g_n} = \frac{(n-k)\frac{p}{q} - (k+1)}{k+1}$$

$$\frac{\Delta g_n}{g_n} = -\frac{(k-n)\frac{p}{q} + k+1}{k+1}$$

$$\frac{\Delta g_n}{g_n} = -\frac{\frac{1}{q}\left((k-n)p + qk\right) + 1}{k+1}$$

$$\frac{\Delta g_n}{g_n} = -\frac{\frac{1}{q}\left(kp-np + qk\right) + 1}{k+1}$$

$$\frac{\Delta g_n}{g_n} = -\frac{\frac{1}{q}\left(k(p+q)-np\right) + 1}{k+1}$$

$$\frac{\Delta g_n}{g_n} = -\frac{\frac{1}{q}\left(k-np\right) + 1}{k+1}$$

$$\frac{\Delta g_n}{g_n} = -\frac{k-np + q}{q(k+1)}$$

$$\frac{\Delta g_n}{g_n} = -\frac{k - np + q}{q(k - np + np + 1)}$$

$$\frac{\Delta g_n}{g_n} = -\frac{\frac{k - np}{\sigma} + q\frac{1}{\sigma}}{q\left(\frac{k - np}{\sigma} + \frac{np}{\sigma} + \frac{1}{\sigma}\right)}$$

$$\frac{\Delta g_n}{g_n} = -\frac{\frac{k - np}{\sigma} + q\frac{1}{\sigma}}{q\frac{k - np}{\sigma} + \frac{npq}{\sigma} + q\frac{1}{\sigma}}$$

Wegen

$$E = np$$

$$\sigma^2 = npq$$

$$\frac{k-np}{\sigma} = \frac{k-E}{\sigma} = x_k$$

$$\Delta x \;=\; x_{k+1} - x_k \;=\; \frac{k+1-np}{\sigma} - \frac{k-np}{\sigma} \;=\; \frac{1}{\sigma}$$

folgt zunächst

$$\frac{\Delta g_n}{g_n} \;=\; -\,\frac{\frac{k\,-\,np}{\sigma} + q\,\frac{1}{\sigma}}{q\,\frac{k\,-\,np}{\sigma} + \frac{npq}{\sigma} + q\,\frac{1}{\sigma}}$$

$$\frac{\Delta g_n}{g_n} \;=\; -\,\frac{x_k + q\,\Delta x}{q\,x_k + \sigma + q\,\Delta x}$$

und weiter

$$\frac{\frac{\Delta g_n}{\Delta x}}{g_n} \;=\; -\,\frac{x_k + q\,\Delta x}{q\,x_k\,\Delta x + \sigma\,\Delta x + q\,\Delta x^2} \quad .$$

Wegen $\sigma\,\Delta x \;=\; 1$ folgt

$$\frac{\frac{\Delta g_n}{\Delta x}}{g_n} \;=\; -\,\frac{x_k + q\,\Delta x}{q\,x_k\,\Delta x + 1 + q\,\Delta x^2} \quad .$$

Betrachte man jetzt die entsprechende Gleichung mit der kontinuierlichen Variablen x, erhält man

$$\frac{\frac{\Delta g_n}{\Delta x}}{g_n} \;=\; -\,\frac{x + q\,\Delta x}{q\,x\,\Delta x + 1 + q\,\Delta x^2} \quad .$$

Für $n \longrightarrow \infty$ nimmt man an , dass gilt :

$$\Delta x \longrightarrow 0$$

$$g_n \longrightarrow \phi$$

$$\frac{\Delta g_n}{\Delta x} \longrightarrow \phi'$$

Damit genügt die Grenzfunktion ϕ der Differentialgleichung

$$\boxed{\frac{\phi'}{\phi} \;=\; -\,x} \quad .$$

Eine Lösung der Differentialgleichung ist :

$$\frac{\phi'}{\phi} = -x$$

$$\frac{1}{\phi}\,\phi' = -x$$

$$\left(\ln\,\phi\right)' = -\left(\frac{1}{2}\,x^2\right)'$$

$$\ln\,\phi = -\frac{1}{2}\,x^2 + C$$

$$\phi = e^{-\frac{1}{2}x^2 + C}$$

$$\underline{\phi = e^C\,e^{-\frac{1}{2}x^2}} \quad .$$

Da die Flächensummen der Funktionen g_n immer gleich 1 sind, muss dies auch für die Grenzfunktion ϕ gelten :

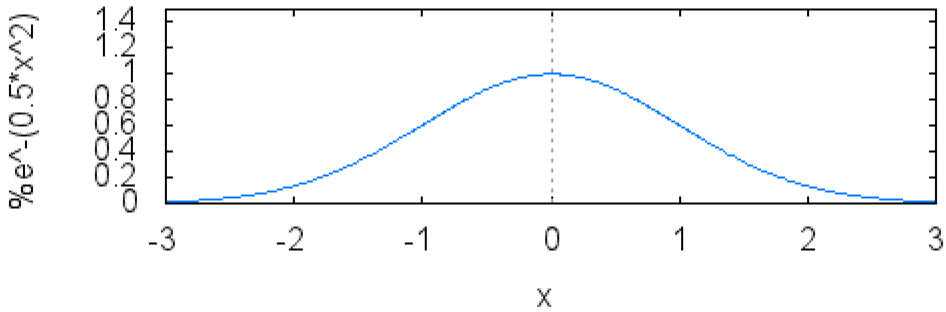
$$\int_{-\infty}^{\infty}\phi(x)\,dx = \int_{-\infty}^{\infty} e^C\,e^{-\frac{1}{2}x^2} = 1$$

Daraus ergibt sich der Wert für die Konstante e^C .

Man zeigt zuerst, dass $\displaystyle\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\,dx = \sqrt{2\pi}$ ist .

wxMaxima 13.04.2

```
wxplot2d([exp(-0.5*x^2) ], [x,-3,3], [y,0,1.5], [gnuplot_preamble, "set size ratio 0.25" ] )$
```



Hierzu betrachtet man den **Rotationskörper**, der entsteht, wenn man den Graphen der Umkehrfunktion von $y = e^{-\frac{1}{2}x^2}$ um die y-Achse rotiert , und bestimmt man das Volumen V .

Die Umkehrfunktion von $y = e^{-\frac{1}{2}x^2}$, I $x \in [0;\infty)$, erhält man durch Umformung:

$$y = e^{-\frac{1}{2}x^2}$$

$$\ln y = -\frac{1}{2}x^2$$

$$-2 \ln y = x^2$$

$$\sqrt{-2 \ln y} = x$$

Die Umkehrfunktion ist $x = \sqrt{-2 \ln y}$, $y \in [0;1]$.

$$V = \int_0^1 \pi x^2 \, dy = \lim_{\epsilon \to 0} \int_\epsilon^1 \pi x^2 \, dy$$

$$\int_\epsilon^1 \pi x^2 \, dy = \int_\epsilon^1 \pi \sqrt{-2 \ln y}^{\,2} \, dy$$

$$\int_\epsilon^1 \pi x^2 \, dy = \int_\epsilon^1 \pi(-2 \ln y) \, dy$$

$$\int_\epsilon^1 \pi x^2 \, dy = -2\pi \int_\epsilon^1 \ln y \, dy$$

Eine Stammfunktion von $\ln y$ ist $y \ln y - y$.

$$\int_\epsilon^1 \pi x^2 \, dy = -2\pi \left[y \ln y - y\right]_\epsilon^1$$

$$\int_\epsilon^1 \pi x^2 \, dy = -2\pi \left[1 \ln 1 - 1 - (\epsilon \ln \epsilon - \epsilon)\right]$$

$$\int_\epsilon^1 \pi x^2 \, dy = -2\pi \left[-1 - (\epsilon \ln \epsilon - \epsilon)\right]$$

$$\int_\epsilon^1 \pi x^2 \, dy = 2\pi \left[1 + (\epsilon \ln \epsilon - \epsilon)\right]$$

$$V = \int_0^1 \pi x^2 \, dy = \lim_{\epsilon \to 0} \int_\epsilon^1 \pi x^2 \, dy$$

$$V = \int_0^1 \pi x^2 \, dy = \lim_{\epsilon \to 0} 2\pi \left[1 + (\epsilon \ln \epsilon - \epsilon)\right]$$

Was ist der Grenzwert $\lim\limits_{\epsilon \to 0} \epsilon \ln \epsilon = \lim\limits_{\epsilon \to 0} \dfrac{\ln \epsilon}{\frac{1}{\epsilon}}$?

Nach der Regel von de **l'Hospital** erhält man :

Guillaume François Antoine, Marquis de L'Hospital oder L'Hôpital, französischer Mathematiker (1661 – 1704)

$$\lim_{\epsilon \to 0} \epsilon \ln \epsilon = \lim_{\epsilon \to 0} \frac{\ln \epsilon}{\frac{1}{\epsilon}}$$

$$\lim_{\epsilon \to 0} \epsilon \ln \epsilon = \lim_{\epsilon \to 0} \frac{\ln' \epsilon}{\left(\frac{1}{\epsilon}\right)'}$$

$$\lim_{\epsilon \to 0} \epsilon \ln \epsilon = \lim_{\epsilon \to 0} \frac{\frac{1}{\epsilon}}{-\frac{1}{\epsilon^2}}$$

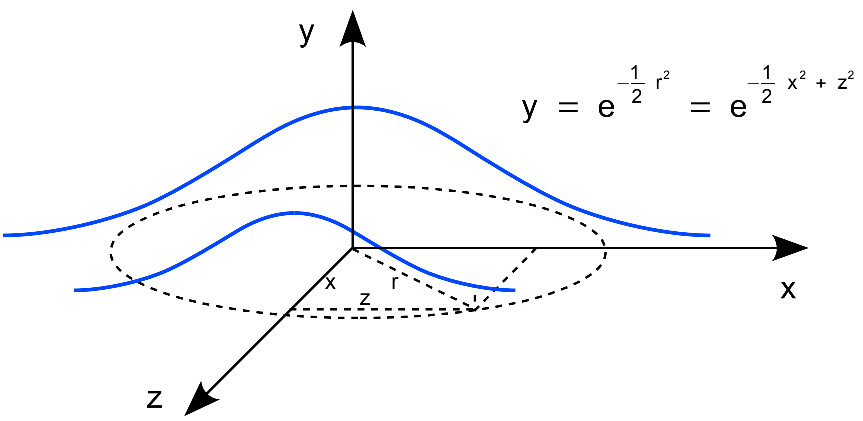
$$\lim_{\epsilon \to 0} \epsilon \ln \epsilon = \lim_{\epsilon \to 0} -\epsilon$$

$$\lim_{\epsilon \to 0} \epsilon \ln \epsilon = 0$$

$$V = \int_0^1 \pi x^2 \, dy = \lim_{\epsilon \to 0} 2\pi \left[1 + (\epsilon \ln \epsilon - \epsilon)\right]$$

$$V = \int_0^1 \pi x^2 \, dy = 2\pi$$

Andererseits kann man das Volumen V des Rotationskörper auch über ein Doppelintegral bestimmen :



Der Rotationskörper ist durch den Graphen der Funktion $y = e^{-\frac{1}{2} x^2 + z^2}$ und die Ebene $y = 0$ begrenzt.

Das Volumen ist :

$$V = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2 + z^2} \, dx \right) dz$$

$$V = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\,z^2} \; e^{-\frac{1}{2}\,x^2} \, dx \right) dz$$

$$V = \int_{-\infty}^{\infty} \left(e^{-\frac{1}{2}\,z^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} \, dx \right) dz$$

$$V = \int_{-\infty}^{\infty} e^{-\frac{1}{2}\,z^2} \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} \, dx \right) dz$$

$$V = \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} \, dx \right) \int_{-\infty}^{\infty} e^{-\frac{1}{2}\,z^2} \, dz$$

$$V = \int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} \, dx \quad \int_{-\infty}^{\infty} e^{-\frac{1}{2}\,z^2} \, dz$$

$$\underline{V = \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} \, dx \right)^2}$$

Jetzt vergleicht man die Volumina und erhält

$$2\pi = \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} \, dx \right)^2 \quad ,$$

$$\underline{\sqrt{2\pi} = \int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} \, dx} \quad .$$

$$\int_{-\infty}^{\infty} \phi(x) \, dx = \int_{-\infty}^{\infty} e^C \, e^{-\frac{1}{2}\,x^2} = 1 \quad .$$

$$\int_{-\infty}^{\infty} e^C \, e^{-\frac{1}{2}\,x^2} = 1$$

$$e^C \int_{-\infty}^{\infty} e^{-\frac{1}{2}\,x^2} = 1$$

$$e^C \sqrt{2\pi} = 1$$

$$e^C = \frac{1}{\sqrt{2\pi}}$$

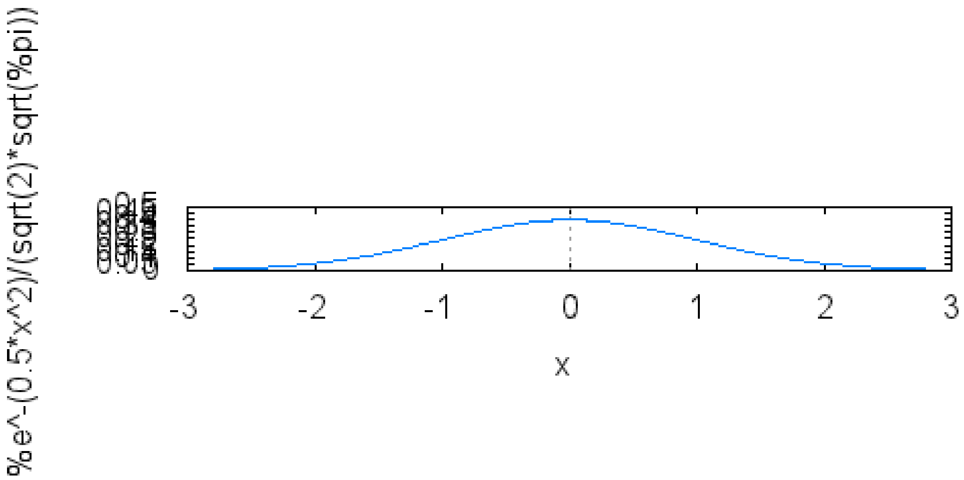
$$\phi(x) = e^C \, e^{-\frac{1}{2}\,x^2}$$

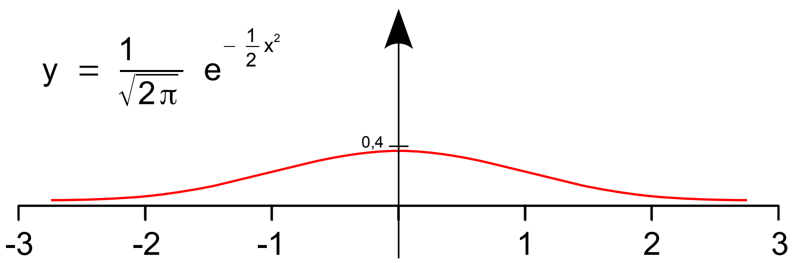
$$\boxed{\phi(x) = \frac{1}{\sqrt{2\pi}} \, e^{-\frac{1}{2}\,x^2}}$$

Die Funktion $\phi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x^2}$ heißt **Gaußfunktion** .

Johann Carl Friedrich Gauß, deutscher Mathematiker (1777 – 1855)

```
wxplot2d([(1/sqrt(2*%pi))*exp(-0.5*x^2) ], [x,-3,3], [y,0,0.5], [gnuplot_preamble, "set size ratio 0.0833" ] )$
```





Wegen $\sigma p_k = g_n\left(\frac{k-E}{\sigma}\right) = g_n(x_k)$ und $g_n \longrightarrow \phi$ für $n \longrightarrow \infty$ gilt

für große n gilt die Näherung :

$$\sigma p_k \approx \phi\left(\frac{k-E}{\sigma}\right)$$

$$p_k \approx \frac{1}{\sigma}\,\phi\left(\frac{k-E}{\sigma}\right)$$

$$p(k) = \binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sigma}\,\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{k-E}{\sigma}\right)^2}$$

Bei der Transformation $x = \frac{k-E}{\sigma}$; $y = \sigma\ p(k)$

werden die Rechteckbreiten durch σ dividiert und die Rechteckhöhen mit σ multipliziert, so dass sich der Flächeninhalt des Histogramms nicht ändert.

Deshalb gilt für die Intervallwahrscheinlichkeit :

$$P(k_1 \le k \le k_2) \approx \int_{\frac{k_1-E}{\sigma}}^{\frac{k_2-E}{\sigma}} \phi(t)\ dt$$

$$P(k_1 \le k \le k_2) \approx \int_{\frac{k_1-E}{\sigma}}^{\frac{k_2-E}{\sigma}} \phi(t)\ dt$$

$$P(k_1 \le k \le k_2) \approx \Phi\left(\frac{k_2-E}{\sigma}\right) - \Phi\left(\frac{k_1-E}{\sigma}\right)$$

Die Funktion Φ heißt **Gaußsche Summenfunktion** und ist definiert als

$$\Phi(x) := \int_{-\infty}^{x} \phi(t)\ dt \quad .$$

Für $\Phi(x)$ lässt sich kein geschlossener Funktionsterm angeben. Ihre Funktionswerte liegen tabelliert vor (siehe nächste Seite) .

Näherungsformel von De Moivre - LaPlace

Abraham de Moivre, französischer Mathematiker (1667 – 1754)
Pierre -Simon (Marquis de) LaPlace, französischer Mathematiker (1749 – 1827)

Für eine $B_{n\,;\,p}$ - verteilte Zufallsvariable gilt für große Werte von n :

$$P(k_1 \le k \le k_2) \approx \Phi\left(\frac{k_2-E}{\sigma}\right) - \Phi\left(\frac{k_1-E}{\sigma}\right) \quad \text{mit} \quad E = np \ , \quad \sigma = \sqrt{np(1-p)} \ .$$

Die Näherung liefert brauchbare Werte, wenn die Faustformel $npq > 9$ erfüllt ist .

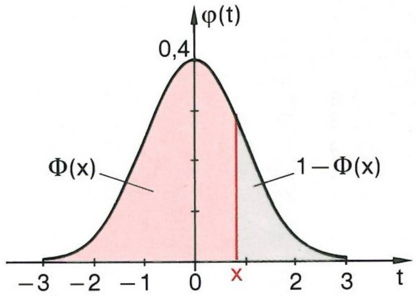
Für $k_1 = 0$ liegt $\Phi\left(\frac{0-E}{\sigma}\right)$ unterhalb der Genauigkeitsgrenze. Damit ergibt sich für

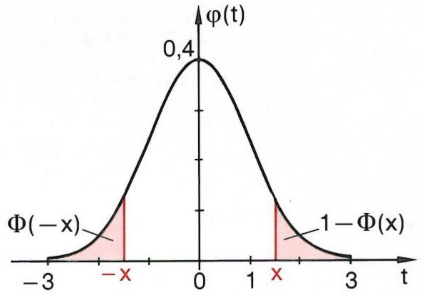
$B_{n\,;\,p}$ - verteilte Zufallsvariablen für große n :

$$P(k \le k) \approx \Phi\left(\frac{k-E}{\sigma}\right) \quad .$$

Tabelle VII: Gaußsche Summenfunktion Φ

$\Phi(x) = 0,...$ $\qquad$ $\Phi(-x) = 1 - \Phi(x)$





x	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	7703	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	9996	9996	9996	9996	9996	9996	9996	9997
3,4	9997	9997	9997	9997	9997	9997	9997	9997	9997	9998

Beispiele: $\Phi(1{,}62) = 0{,}9474$, $\qquad$ $\Phi(-1{,}62) = 1 - 0{,}9474 = 0{,}0526$,

$\Phi(x) = 0{,}6772 \Rightarrow x = 0{,}46$, $\quad$ $\Phi(x) = 0{,}3228 = 1 - 0{,}6772 \Rightarrow x = -0{,}46$

Einige besondere Werte:

x	1,2816	1,6449	1,9600	2,3263	2,5758	3,0902	3,2905
$\Phi(x)$	0,9	0,95	0,975	0,99	0,995	0,999	0,9995

269

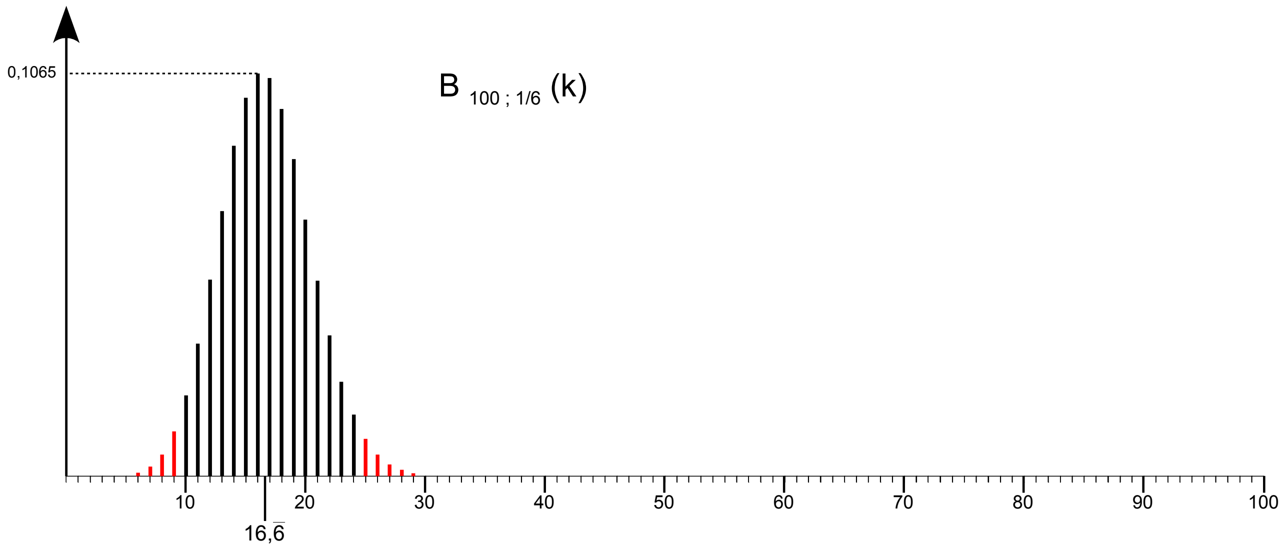
Testprobleme

Zweiseitiger Test

1. Beispiel

Gegeben ist ein Würfel, von dem man nicht weiß, wie groß die Wahrscheinlichkeit p ist, eine Eins zu werfen. Die Vermutung oder **Hypothese** ist $p = \frac{1}{6}$. Die Hypothese soll durch Experiment und durch folgende Überlegungen bestätigt oder abgelehnt werden .

Der Würfel wird zum Beispiel 100-mal geworfen. Die Wahrscheinlichkeitsverteilung $B_{100;\frac{1}{6}}(k)$ hat den **Erwartungswert** $E = 100 \cdot \frac{1}{6} = 16,\overline{6}$.



Die Wahrscheinlichkeit , dass k ein „Ausrutscher“ ist, also etwa $k < 10$ oder $24 < k$, kann folgendermaßen berechnet werden :

$$p(k < 10) = \sum_{k=0}^{9} \binom{100}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}$$

$$p(k < 10) = F_{100;\frac{1}{6}}(9)$$

$$p(k < 10) = 0{,}0213$$

$$p(k > 24) = 1 - P(k \leq 24)$$

$$p(k > 24) = 1 - \sum_{k=0}^{24} \binom{100}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}$$

$$p(k > 24) = 1 - F_{100;\frac{1}{6}}(24)$$

$$p(k > 24) = 1 - 0{,}9783$$

$$p(k > 24) = 0{,}0217$$

Die numerischen Werte für $F_{100;\frac{1}{6}}(k)$ sind der Tabelle für die **Summenfunktion der Binomialverteilung** $B_{100;\frac{1}{6}}(k)$ entnommen (siehe nächste Seite) .

$$p(k < 10 \quad \text{oder} \quad k > 24) = p(k < 10) + p(k > 24)$$

$$p(k < 10 \quad \text{oder} \quad k > 24) = 0{,}0213 + 0{,}0217$$

$$p(k < 10 \quad \text{oder} \quad k > 24) = 0{,}0430$$

$$p(k < 10 \quad \text{oder} \quad k > 24) < 0{,}05$$

$$p(k < 10 \quad \text{oder} \quad k > 24) < 5\%$$

$$p(10 \leq k \leq 24) = 1 - p(k < 10 \quad \text{oder} \quad k > 24)$$

$$p(10 \leq k \leq 24) > 1 - 5\%$$

$$p(10 \leq k \leq 24) > 95\%$$

Die Wahrscheinlichkeit, dass k im „Trefferbereich“ zwischen 10 und 24 einschließlich der Grenzen liegt, ist höher als 95% .

Wenn nun beim Experiment k -mal die 1 geworfen wird mit $k \in [10;24]$, wird man die Hypothese $p(1) = \frac{1}{6}$ annehmen.

Wird aber beim Experiment k -mal die 1 geworfen wird mit $k \in [0;9] \cup [24;100]$, wird man die Hypothese $p = \frac{1}{6}$ ablehnen . Der Bereich $[0;9] \cup [24;100]$ heißt **Ablehnungsbereich** .

Die Wahrscheinlichkeit, dass man bei dieser Entscheidung einen Fehler begeht, dass also das Ergebnis in Wahrheit ein statistischer „Ausrutscher“ ist, beträgt weniger als 5% . Diese Wahrscheinlichkeit heißt **Irrtumswahrscheinlichkeit** .

Bemerkung :
Beim Testen einer Wahrscheinlichkeit für Treffer, wird die Irrtumswahrscheinlichkeit immer vor dem Experiment festgelegt .

Tabelle V: Binomialverteilung

$F_{n;p}(k) = B_{n;p}(0) + \ldots + B_{n;p}(k)$

Summenfunktion

n	k	0,02	0,03	0,04	0,05	0,10	1/6	0,20	0,30	1/3	0,40	0,50	k	n
	0	0,1326	0476	0169	0059	0000	0000	0000	0000	0000	0000	0000	99	
	1	4033	1946	0872	0371	0003	0000	0000	0000	0000	0000	0000	98	
	2	6767	4198	2321	1183	0019	0000	0000	0000	0000	0000	0000	97	
	3	8590	6472	4295	2578	0078	0000	0000	0000	0000	0000	0000	96	
	4	9492	8179	6289	4360	0237	0001	0000	0000	0000	0000	0000	95	
	5	9845	9192	7884	6160	0576	0004	0000	0000	0000	0000	0000	94	
	6	9959	9688	8936	7660	1172	0013	0001	0000	0000	0000	0000	93	
	7	9991	9894	9525	8720	2061	0038	0003	0000	0000	0000	0000	92	
	8	9998	9968	9810	9369	3209	0095	0009	0000	0000	0000	0000	91	
	9		9991	9932	9718	4513	0213	0023	0000	0000	0000	0000	90	
	10		9998	9978	9885	5832	0427	0057	0000	0000	0000	0000	89	
	11			9993	9957	7030	0777	0126	0000	0000	0000	0000	88	
	12			9998	9985	8018	1297	0253	0000	0000	0000	0000	87	
	13				9995	8761	2000	0469	0001	0000	0000	0000	86	
	14				9999	9274	2874	0804	0002	0000	0000	0000	85	
	15					9601	3877	1285	0004	0000	0000	0000	84	
	16					9794	4942	1923	0010	0001	0000	0000	83	
	17					9900	5994	2712	0022	0002	0000	0000	82	
	18					9954	6965	3621	0045	0005	0000	0000	81	
	19					9980	7803	4602	0089	0011	0000	0000	80	
	20					9992	8481	5595	0165	0024	0000	0000	79	
	21					9997	8998	6540	0288	0048	0000	0000	78	
	22					9999	9370	7389	0479	0091	0001	0000	77	
	23						9621	8109	0755	0164	0003	0000	76	
	24						9783	8686	1136	0281	0006	0000	75	
	25						9881	9125	1631	0458	0012	0000	74	
	26						9938	9442	2244	0715	0024	0000	73	
	27						9969	9658	2964	1066	0046	0000	72	
	28						9985	9800	3768	1524	0084	0000	71	
	29						9993	9888	4623	2093	0148	0000	70	
	30						9997	9939	5491	2766	0248	0000	69	
	31						9999	9969	6331	3525	0398	0001	68	
	32							9985	7107	4344	0615	0002	67	
	33							9993	7793	5188	0913	0004	66	
	34							9997	8371	6019	1303	0009	65	
	35							9999	8839	6803	1795	0018	64	
100	36							9999	9201	7511	2386	0033	63	100
	37								9470	8123	3068	0060	62	
	38								9660	8630	3822	0105	61	
	39								9790	9034	4621	0176	60	
	40								9875	9341	5433	0284	59	
	41								9928	9566	6225	0443	58	
	42								9960	9724	6967	0666	57	
	43								9979	9831	7635	0967	56	
	44								9989	9900	8211	1356	55	
	45								9995	9943	8689	1841	54	
	46								9997	9969	9070	2421	53	
	47								9999	9983	9362	3087	52	
	48								9999	9991	9577	3822	51	
	49									9996	9729	4602	50	
	50									9998	9832	5398	49	
	51									9999	9900	6178	48	
	52										9942	6914	47	
	53										9968	7579	46	
	54										9983	8159	45	
	55										9991	8644	44	
	56										9996	9033	43	
	57										9998	9334	42	
	58										9999	9557	41	
	59											9716	40	
	60											9824	39	
	61											9895	38	
	62											9940	37	
	63											9967	36	
	64											9982	35	
	65											9991	34	
	66											9996	33	
	67											9998	32	
	68											9999	31	
n		0,98	0,97	0,96	0,95	0,90	5/6	0,80	0,70	2/3	0,60	0,50	k	n

p

Bei rot gedrucktem Eingang, d. h. $p \geqq 0{,}5$ gilt:
$F_{n;p}(k) = 1 -$ abgelesener Wert

Nicht aufgeführte Werte sind (auf 4 Dez.) 1,0000.

Testprobleme

Zweiseitiger Test

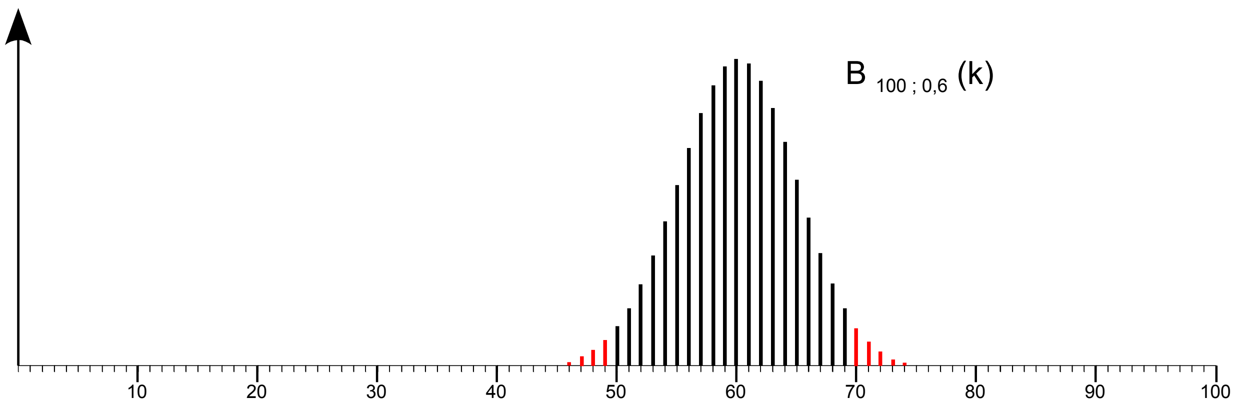
2. Beispiel

Der Bürgermeister einer Stadt erhielt bei der letzten Wahl $60\,\%$ der Stimmen. Bei der Befragung von $n = 100$ Personen vor der kommenden Wahl bevorzugten $k = 48$ Personen den Bürgermeister.
Kann man mit einer **Irrtumswahrscheinlichkeit** $\alpha = 5\%$ schließen, dass sich der Stimmenanteil verändert hat ?

Hypothese H_0 : $p = 60\,\% = 0{,}6$.

Gegenhypothese H_1 : $p \neq 0{,}6$.

Bestimmung des **Ablehnungsbereichs** mit Hilfe der Tabelle für die **Summenfunktion der Binomialverteilung** $F_{100\,;\,0{,}6}(k)$ (siehe nächste Seite) .



$$p(k \leq k_1) \leq 2{,}5\%$$
$$p(k \leq k_1) \leq 0{,}025$$
$$k_1 = 49$$

$$p(k_2 \leq k) \leq 2{,}5\%$$
$$p(k_2 \leq k) \leq 0{,}025$$
$$1 - p(k < k_2) \leq 0{,}025$$
$$1 - p(k \leq k_2 - 1) \leq 0{,}025$$
$$p(k \leq k_2 - 1) \geq 1 - 0{,}025$$
$$p(k \leq k_2 - 1) \geq 0{,}0975$$
$$k_2 - 1 = 69$$
$$k_2 = 70$$

Ablehnungsbereich : $A = [0;49] \cup [70;100]$.

Wegen $k = 48 \in A$ wird die **Hypothese** H_0 , dass $p = 60\,\%$ ist, **abgelehnt** .
Die Wahrscheinlichkeit, dass man bei dieser Entscheidung einen Fehler begeht, dass also das Ergebnis in Wahrheit ein statistischer „Ausrutscher“ ist, beträgt höchstens $5\,\%$.

Tabelle V: Binomialverteilung

Summenfunktion

$F_{n;p}(k) = B_{n;p}(0) + \ldots + B_{n;p}(k)$

n	k	0,02	0,03	0,04	0,05	0,10	1/6	0,20	0,30	1/3	0,40	0,50	k	n
	0	0,1326	0476	0169	0059	0000	0000	0000	0000	0000	0000	0000	99	
	1	4033	1946	0872	0371	0003	0000	0000	0000	0000	0000	0000	98	
	2	6767	4198	2321	1183	0019	0000	0000	0000	0000	0000	0000	97	
	3	8590	6472	4295	2578	0078	0000	0000	0000	0000	0000	0000	96	
	4	9492	8179	6289	4360	0237	0001	0000	0000	0000	0000	0000	95	
	5	9845	9192	7884	6160	0576	0004	0000	0000	0000	0000	0000	94	
	6	9959	9688	8936	7660	1172	0013	0001	0000	0000	0000	0000	93	
	7	9991	9894	9525	8720	2061	0038	0003	0000	0000	0000	0000	92	
	8	9998	9968	9810	9369	3209	0095	0009	0000	0000	0000	0000	91	
	9		9991	9932	9718	4513	0213	0023	0000	0000	0000	0000	90	
	10		9998	9978	9885	5832	0427	0057	0000	0000	0000	0000	89	
	11			9993	9957	7030	0777	0126	0000	0000	0000	0000	88	
	12			9998	9985	8018	1297	0253	0000	0000	0000	0000	87	
	13				9995	8761	2000	0469	0001	0000	0000	0000	86	
	14				9999	9274	2874	0804	0002	0000	0000	0000	85	
	15					9601	3877	1285	0004	0000	0000	0000	84	
	16					9794	4942	1923	0010	0001	0000	0000	83	
	17					9900	5994	2712	0022	0002	0000	0000	82	
	18					9954	6965	3621	0045	0005	0000	0000	81	
	19					9980	7803	4602	0089	0011	0000	0000	80	
	20					9992	8481	5595	0165	0024	0000	0000	79	
	21					9997	8998	6540	0288	0048	0000	0000	78	
	22					9999	9370	7389	0479	0091	0001	0000	77	
	23						9621	8109	0755	0164	0003	0000	76	
	24						9783	8686	1136	0281	0006	0000	75	
	25						9881	9125	1631	0458	0012	0000	74	
	26						9938	9442	2244	0715	0024	0000	73	
	27						9969	9658	2964	1066	0046	0000	72	
	28						9985	9800	3768	1524	0084	0000	71	
	29						9993	9888	4623	2093	0148	0000	70	
	30						9997	9939	5491	2766	0248	0000	69	
	31						9999	9969	6331	3525	0398	0001	68	
	32							9985	7107	4344	0615	0002	67	
	33							9993	7793	5188	0913	0004	66	
	34							9997	8371	6019	1303	0009	65	
	35							9999	8839	6803	1795	0018	64	
100	36							9999	9201	7511	2386	0033	63	100
	37								9470	8123	3068	0060	62	
	38								9660	8630	3822	0105	61	
	39								9790	9034	4621	0176	60	
	40								9875	9341	5433	0284	59	
	41								9928	9566	6225	0443	58	
	42								9960	9724	6967	0666	57	
	43								9979	9831	7635	0967	56	
	44								9989	9900	8211	1356	55	
	45								9995	9943	8689	1841	54	
	46								9997	9969	9070	2421	53	
	47								9999	9983	9362	3087	52	
	48								9999	9991	9577	3822	51	
	49									9996	9729	4602	50	
	50									9998	9832	5398	49	
	51									9999	9900	6178	48	
	52										9942	6914	47	
	53										9968	7579	46	
	54										9983	8159	45	
	55										9991	8644	44	
	56										9996	9033	43	
	57										9998	9334	42	
	58										9999	9557	41	
	59											9716	40	
	60											9824	39	
	61											9895	38	
	62											9940	37	
	63											9967	36	
	64											9982	35	
	65											9991	34	
	66											9996	33	
	67											9998	32	
	68											9999	31	
n		0,98	0,97	0,96	0,95	0,90	5/6	0,80	0,70	2/3	0,60	0,50	k	n

p

Bei rot gedrucktem Eingang, d. h. $p \geqq 0{,}5$ gilt:
$F_{n;p}(k) = 1 -$ abgelesener Wert

Nicht aufgeführte Werte sind (auf 4 Dez.) 1,0000.

Fehler 1. Art und Fehler 2. Art

Die folgenden Überlegungen beziehen sich zahlenmäßig auf das **2. Beispiel** , auf die Bürgermeisterwahl.

Man macht einen **Fehler 1. Art**, wenn man die Hypothese H_0 , dass $p = 0{,}6$ ist, ablehnt, obwohl diese richtig ist.

Die **Irrtumswahrscheinlichkeit** $\alpha = 5\%$ ist die Wahrscheinlichkeit, die Hypothese H_0 abzulehnen, obwohl diese richtig ist .

Man macht einen **Fehler 2. Art**, wenn man die Hypothese H_0 , dass $p = 0{,}6$ ist, beibehält, obwohl diese falsch ist.

Die **Irrtumswahrscheinlichkeit** β ist die Wahrscheinlichkeit, die Hypothese H_0 beizubehalten, obwohl diese falsch ist. Diese Irrtumswahrscheinlichkeit kann man nicht angeben, da es unendlich viele $p \neq 60\%$ gibt .
Nimmt man aber zum Beispiel an, dass in Wirklichkeit $p = 0{,}5$ ist, so folgt bei dem vorgegebenen Ablehnungsbereich $A = [0;49] \cup [70;100]$:

$$\beta = \sum_{k=50}^{69} B_{100;0,5}(k)$$

$$\beta = F_{100;0,5}(69) - F_{100;0,5}(50-1)$$

$$\beta = F_{100;0,5}(69) - F_{100;0,5}(49)$$

$$\beta \approx 0{,}9999 - 0{,}4602$$

$$\beta \approx 0{,}5379$$

$$\beta \approx 54\%$$

Die Irrtumswahrscheinlichkeit β beim Fehler 2. Art kann also recht hoch sein im Vergleich zur Irrtumswahrscheinlichkeit α beim Fehler 1. Art .

Testprobleme

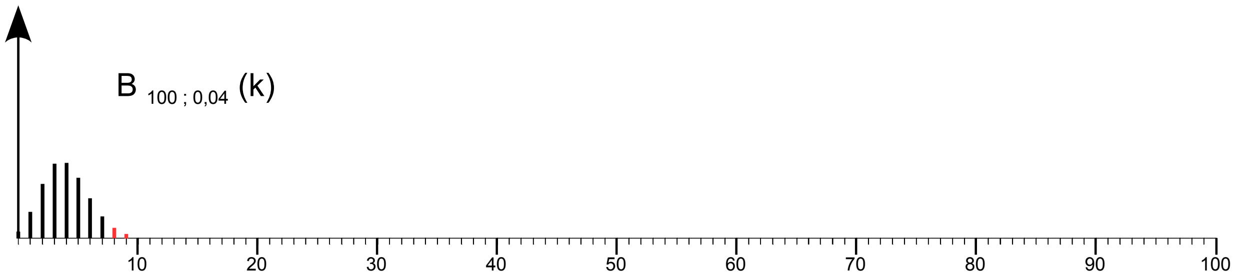
Rechtsseitiger Test

Der Hersteller eines Artikels garantiert, dass höchstens $4\,\%$ Makulatur sind .
Es werden $n = 100$ Artikel einer Lieferung entnommen, und man findet $k = 9$ fehlerhafte Artikel .
Kann man mit der Irrtumswahrscheinlichkeit $\alpha = 5\%$ schließen, dass die Makulatur höher als $4\,\%$ ist ?

Hypothese H_0 : $p \leq 0{,}04$.

Gegenhypothese H_1 : $p > 0{,}04$.

Bestimmung des **Ablehnungsbereichs** mit Hilfe der Tabelle für die **Summenfunktion der Binomialverteilung** $F_{100\,;\,0{,}04}(k)$ (siehe nächste Seite) .



$$p(k_2 \leq k) \leq 5\,\%$$
$$p(k_2 \leq k) \leq 0{,}05$$
$$1 - p(k < k_2) \leq 0{,}05$$
$$1 - p(k \leq k_2 - 1) \leq 0{,}05$$
$$p(k \leq k_2 - 1) \geq 1 - 0{,}05$$
$$p(k \leq k_2 - 1) \geq 0{,}95$$
$$F_{100;0{,}04}(k_2 - 1) \geq 0{,}95$$
$$F_{100;0{,}04}(6) \approx 0{,}8936$$
$$F_{100;0{,}04}(7) \approx 0{,}9525$$
$$k_2 - 1 = 7$$
$$k_2 = 8$$

Ablehnungsbereich : $A = [8;100]$.

Da $k = 9 \in A = [8;100]$ wird die Hypothese H_0 , dass $p \leq 0{,}04$ ist, abgelehnt .

Mit der Irrtumswahrscheinlichkeit $\alpha = 5\%$ kann man sagen, dass die Gegenhypothese gilt, dass $p > 0{,}04 = 4\,\%$ ist .

Tabelle V: Binomialverteilung

$F_{n;p}(k) = B_{n;p}(0) + \ldots + B_{n;p}(k)$

Summenfunktion

n	k	p = 0,02	0,03	0,04	0,05	0,10	1/6	0,20	0,30	1/3	0,40	0,50	k	n
	0	0,1326	0476	0169	0059	0000	0000	0000	0000	0000	0000	0000	99	
	1	4033	1946	0872	0371	0003	0000	0000	0000	0000	0000	0000	98	
	2	6767	4198	2321	1183	0019	0000	0000	0000	0000	0000	0000	97	
	3	8590	6472	4295	2578	0078	0000	0000	0000	0000	0000	0000	96	
	4	9492	8179	6289	4360	0237	0001	0000	0000	0000	0000	0000	95	
	5	9845	9192	7884	6160	0576	0004	0000	0000	0000	0000	0000	94	
	6	9959	9688	8936	7660	1172	0013	0001	0000	0000	0000	0000	93	
	7	9991	9894	9525	8720	2061	0038	0003	0000	0000	0000	0000	92	
	8	9998	9968	9810	9369	3209	0095	0009	0000	0000	0000	0000	91	
	9		9991	9932	9718	4513	0213	0023	0000	0000	0000	0000	90	
	10		9998	9978	9885	5832	0427	0057	0000	0000	0000	0000	89	
	11			9993	9957	7030	0777	0126	0000	0000	0000	0000	88	
	12			9998	9985	8018	1297	0253	0000	0000	0000	0000	87	
	13				9995	8761	2000	0469	0001	0000	0000	0000	86	
	14				9999	9274	2874	0804	0002	0000	0000	0000	85	
	15					9601	3877	1285	0004	0000	0000	0000	84	
	16					9794	4942	1923	0010	0001	0000	0000	83	
	17					9900	5994	2712	0022	0002	0000	0000	82	
	18					9954	6965	3621	0045	0005	0000	0000	81	
	19					9980	7803	4602	0089	0011	0000	0000	80	
	20					9992	8481	5595	0165	0024	0000	0000	79	
	21					9997	8998	6540	0288	0048	0000	0000	78	
	22					9999	9370	7389	0479	0091	0001	0000	77	
	23						9621	8109	0755	0164	0003	0000	76	
	24						9783	8686	1136	0281	0006	0000	75	
	25						9881	9125	1631	0458	0012	0000	74	
	26						9938	9442	2244	0715	0024	0000	73	
	27						9969	9658	2964	1066	0046	0000	72	
	28						9985	9800	3768	1524	0084	0000	71	
	29						9993	9888	4623	2093	0148	0000	70	
	30						9997	9939	5491	2766	0248	0000	69	
	31						9999	9969	6331	3525	0398	0001	68	
	32							9985	7107	4344	0615	0002	67	
	33							9993	7793	5188	0913	0004	66	
	34							9997	8371	6019	1303	0009	65	
	35							9999	8839	6803	1795	0018	64	
100	36							9999	9201	7511	2386	0033	63	100
	37								9470	8123	3068	0060	62	
	38								9660	8630	3822	0105	61	
	39								9790	9034	4621	0176	60	
	40								9875	9341	5433	0284	59	
	41								9928	9566	6225	0443	58	
	42								9960	9724	6967	0666	57	
	43								9979	9831	7635	0967	56	
	44								9989	9900	8211	1356	55	
	45								9995	9943	8689	1841	54	
	46								9997	9969	9070	2421	53	
	47								9999	9983	9362	3087	52	
	48								9999	9991	9577	3822	51	
	49									9996	9729	4602	50	
	50									9998	9832	5398	49	
	51									9999	9900	6178	48	
	52										9942	6914	47	
	53										9968	7579	46	
	54										9983	8159	45	
	55										9991	8644	44	
	56										9996	9033	43	
	57										9998	9334	42	
	58										9999	9557	41	
	59											9716	40	
	60											9824	39	
	61											9895	38	
	62											9940	37	
	63											9967	36	
	64											9982	35	
	65											9991	34	
	66											9996	33	
	67											9998	32	
	68											9999	31	
n		0,98	0,97	0,96	0,95	0,90	5/6	0,80	0,70	2/3	0,60	0,50	k	n
							p							

Bei rot gedrucktem Eingang, d. h. $p \geqq 0{,}5$ gilt:
$F_{n;p}(k) = 1 -$ abgelesener Wert

Nicht aufgeführte Werte sind (auf 4 Dez.) 1,0000.

Testprobleme

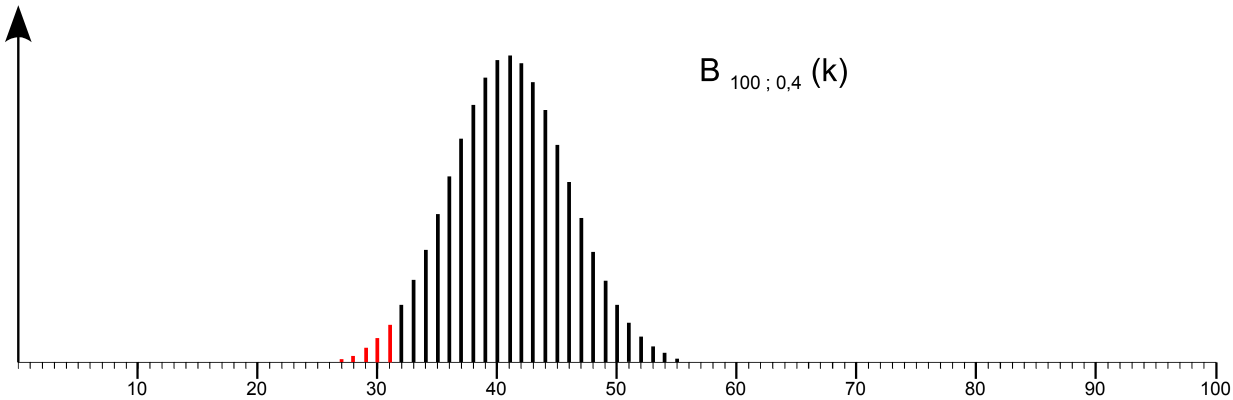
Linksseitiger Test

Ein Kandidat hatte bei der letzten Wahl $40\,\%$ der Stimmen . Vor der nächsten Wahl wird eine Umfrage gestartet . Von $n = 100$ Personen geben $k = 34$ Personen an, den Kandidaten wählen zu wollen .
Kann man mit der Irrtumswahrscheinlichkeit $\alpha = 5\%$ schließen, dass der Stimmenanteil gesunken ist ?

Hypothese H_0 : $p \geq 40\% = 0{,}4$.

Gegenhypothese H_1 : $p < 0{,}4$.

Bestimmung des **Ablehnungsbereichs** mit Hilfe der Tabelle für die **Summenfunktion der Binomialverteilung** $F_{100\,;\,0{,}4}(k)$ (siehe nächste Seite) .



$p(k \leq k_1) \leq 5\%$
$p(k \leq k_1) \leq 0{,}05$
$F_{100\,;\,0{,}4}(k_1) \leq 0{,}05$
$F_{100\,;\,0{,}4}(31) \approx 0{,}0398$
$F_{100\,;\,0{,}4}(32) \approx 0{,}0615$
$k_1 = 31$

Ablehnungsbereich : $A = [0;31]$.

Da $k = 34 \notin A = [0;31]$ wird die Hypothese H_0 , dass $p \geq 0{,}4$ ist, akzeptiert .

Mit der Irrtumswahrscheinlichkeit $\alpha = 5\%$ kann man davon ausgehen, dass der Kandidat keine Stimmen verlieren wird, dass also $p \geq 0{,}4 = 40\,\%$ ist .

Tabelle V: Binomialverteilung

Summenfunktion

$F_{n;p}(k) = B_{n;p}(0) + \ldots + B_{n;p}(k)$

n	k	0,02	0,03	0,04	0,05	0,10	1/6	0,20	0,30	1/3	0,40	0,50		n
	0	0,1326	0476	0169	0059	0000	0000	0000	0000	0000	0000	0000	99	
	1	4033	1946	0872	0371	0003	0000	0000	0000	0000	0000	0000	98	
	2	6767	4198	2321	1183	0019	0000	0000	0000	0000	0000	0000	97	
	3	8590	6472	4295	2578	0078	0000	0000	0000	0000	0000	0000	96	
	4	9492	8179	6289	4360	0237	0001	0000	0000	0000	0000	0000	95	
	5	9845	9192	7884	6160	0576	0004	0000	0000	0000	0000	0000	94	
	6	9959	9688	8936	7660	1172	0013	0001	0000	0000	0000	0000	93	
	7	9991	9894	9525	8720	2061	0038	0003	0000	0000	0000	0000	92	
	8	9998	9968	9810	9369	3209	0095	0009	0000	0000	0000	0000	91	
	9		9991	9932	9718	4513	0213	0023	0000	0000	0000	0000	90	
	10		9998	9978	9885	5832	0427	0057	0000	0000	0000	0000	89	
	11			9993	9957	7030	0777	0126	0000	0000	0000	0000	88	
	12			9998	9985	8018	1297	0253	0000	0000	0000	0000	87	
	13				9995	8761	2000	0469	0001	0000	0000	0000	86	
	14				9999	9274	2874	0804	0002	0000	0000	0000	85	
	15					9601	3877	1285	0004	0000	0000	0000	84	
	16					9794	4942	1923	0010	0001	0000	0000	83	
	17					9900	5994	2712	0022	0002	0000	0000	82	
	18					9954	6965	3621	0045	0005	0000	0000	81	
	19					9980	7803	4602	0089	0011	0000	0000	80	
	20					9992	8481	5595	0165	0024	0000	0000	79	
	21					9997	8998	6540	0288	0048	0000	0000	78	
	22					9999	9370	7389	0479	0091	0001	0000	77	
	23						9621	8109	0755	0164	0003	0000	76	
	24						9783	8686	1136	0281	0006	0000	75	
	25						9881	9125	1631	0458	0012	0000	74	
	26						9938	9442	2244	0715	0024	0000	73	
	27						9969	9658	2964	1066	0046	0000	72	
	28						9985	9800	3768	1524	0084	0000	71	
	29						9993	9888	4623	2093	0148	0000	70	
	30						9997	9939	5491	2766	0248	0000	69	
	31						9999	9969	6331	3525	0398	0001	68	
	32							9985	7107	4344	0615	0002	67	
	33							9993	7793	5188	0913	0004	66	
	34							9997	8371	6019	1303	0009	65	
	35							9999	8839	6803	1795	0018	64	
100	36							9999	9201	7511	2386	0033	63	100
	37								9470	8123	3068	0060	62	
	38								9660	8630	3822	0105	61	
	39								9790	9034	4621	0176	60	
	40								9875	9341	5433	0284	59	
	41								9928	9566	6225	0443	58	
	42								9960	9724	6967	0666	57	
	43								9979	9831	7635	0967	56	
	44								9989	9900	8211	1356	55	
	45								9995	9943	8689	1841	54	
	46								9997	9969	9070	2421	53	
	47								9999	9983	9362	3087	52	
	48								9999	9991	9577	3822	51	
	49									9996	9729	4602	50	
	50									9998	9832	5398	49	
	51									9999	9900	6178	48	
	52										9942	6914	47	
	53										9968	7579	46	
	54										9983	8159	45	
	55										9991	8644	44	
	56										9996	9033	43	
	57										9998	9334	42	
	58										9999	9557	41	
	59											9716	40	
	60											9824	39	
	61											9895	38	
	62											9940	37	
	63											9967	36	
	64											9982	35	
	65											9991	34	
	66											9996	33	
	67											9998	32	
	68											9999	31	
n		0,98	0,97	0,96	0,95	0,90	5/6	0,80	0,70	2/3	0,60	0,50	k	n

p

Bei rot gedrucktem Eingang, d. h. $p \geqq 0{,}5$ gilt:
$F_{n;p}(k) = 1 -$ abgelesener Wert

Nicht aufgeführte Werte sind (auf 4 Dez.) 1,0000.

Schätzprobleme

Im Gegensatz zu Testproblemen bei denen man immer eine Vermutung über den zu testenden Parameter hat, hat man bei Schätzproblemen praktisch keine Anhaltspunkte .

Beispiel

Aus einer Urne mit sehr vielen schwarzen und weißen Kugeln wird eine Stichprobe mit großem Umfang n entnommen . Es wird nach der Wahrscheinlichkeit p für eine schwarze Kugel gefragt .

Hierzu betrachtet man die **relative Häufigkeit** $\bar{k} = \frac{k}{n}$ für das Vorkommen einer schwarzen Kugel , die nach dem Gesetz der großen Zahlen immer weniger von p abweicht .

Die Variable k ist $B_{n\,;\,p}$ - verteilt und werde durch die **Gaußfunktion** φ approximiert . Das heißt :

$$p(k) = b_{n:p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(k_1 \le k \le k_2) = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(k_1 \le k \le k_2) \approx \int_{\frac{k_1-E}{\sigma}}^{\frac{k_2-E}{\sigma}} \phi(t)\ dt$$

mit $\phi(t) := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$, $E = np$, $\sigma = \sqrt{np(1-p)}$

$$P(k_1 \le k \le k_2) \approx \int_{\frac{k_1-E}{\sigma}}^{\frac{k_2-E}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\ dt$$

$$P(k_1 \le k \le k_2) \approx \Phi\left(\frac{k_2-E}{\sigma}\right) - \Phi\left(\frac{k_1-E}{\sigma}\right) \quad \text{mit} \quad \Phi(x) := \int_{-\infty}^{x} \phi(t)\ dt$$

Die Wahrscheinlichkeit, dass $\frac{k}{n}$ im Intervall $\left[p - \frac{c\sigma}{n}\ ;\ p + \frac{c\sigma}{n}\right]$, also im Intervall um p der Länge $2\,\frac{c\sigma}{n}$ liegt, ist gegeben durch :

$$P\left(\left|\frac{k}{n} - p\right| \le \frac{c\sigma}{n}\right) = P\left(p - \frac{c\sigma}{n} \le \frac{k}{n} \le p + \frac{c\sigma}{n}\right)$$

$$P\left(\left|\frac{k}{n} - p\right| \le \frac{c\,\sigma}{n}\right) = P(np - c\sigma \le k \le np + c\sigma)$$

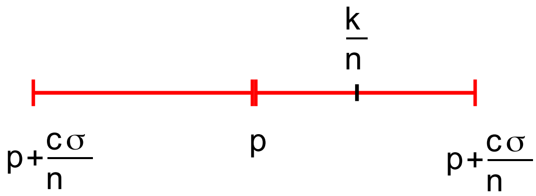
$$P\left(\left|\frac{k}{n} - p\right| \le \frac{c\,\sigma}{n}\right) = P(E - c\sigma \le k \le E + c\sigma)$$

$$P\left(\left|\frac{k}{n} - p\right| \le \frac{c\,\sigma}{n}\right) \approx \Phi\left(\frac{E+c\sigma - E}{\sigma}\right) - \Phi\left(\frac{E-c\sigma - E}{\sigma}\right)$$

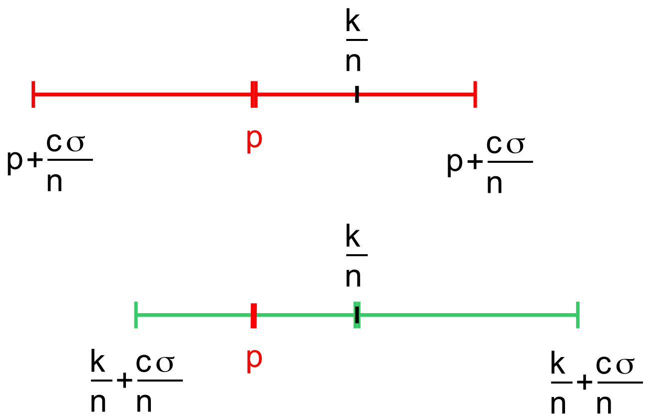
$$P\left(\left|\frac{k}{n} - p\right| \le \frac{c\,\sigma}{n}\right) \approx \Phi(c) - \Phi(-c)$$

$$P\left(\left|\frac{k}{n} - p\right| \le \frac{c\,\sigma}{n}\right) \approx \Phi(c) - (1 - \Phi(c))$$

$$P\left(\left|\frac{k}{n} - p\right| \le \frac{c\,\sigma}{n}\right) \approx 2\,\Phi(c) - 1$$



Zeigt das Experiment jetzt eine zufällige relative Häufigkeit $\frac{k}{n}$, so denkt man sich um $\frac{k}{n}$ ebenfalls ein Intervall der Länge $2\,\frac{c\sigma}{n}$:



Die Wahrscheinlichkeit , dass p in diesem zufälligen Intervall um $\frac{k}{n}$ liegt ist :

$$P\left(\left|p - \frac{k}{n}\right| \le \frac{c\,\sigma}{n}\right) = P\left(\frac{k}{n} - \frac{c\,\sigma}{n} \le p \le \frac{k}{n} + \frac{c\,\sigma}{n}\right)$$

$$P\left(\left|p - \frac{k}{n}\right| \le \frac{c\,\sigma}{n}\right) = P\left(-\frac{k}{n} - \frac{c\,\sigma}{n} \le -p \le -\frac{k}{n} + \frac{c\,\sigma}{n}\right)$$

$$P\left(\left|p - \frac{k}{n}\right| \le \frac{c\,\sigma}{n}\right) = P\left(-\frac{c\,\sigma}{n} \le \frac{k}{n} - p \le \frac{c\,\sigma}{n}\right)$$

$$P\left(\left|p - \frac{k}{n}\right| \le \frac{c\,\sigma}{n}\right) = P\left(p - \frac{c\,\sigma}{n} \le \frac{k}{n} \le p + \frac{c\,\sigma}{n}\right)$$

$$P\left(\left|p - \frac{k}{n}\right| \le \frac{c\,\sigma}{n}\right) = P\left(\left|\frac{k}{n} - p\right| \le \frac{c\,\sigma}{n}\right)$$

$$P\left(\left|p - \frac{k}{n}\right| \le \frac{c\,\sigma}{n}\right) \approx 2\ \Phi(c) - 1$$

Die Wahrscheinlichkeit , dass p in diesem zufälligen Intervall um $\frac{k}{n}$ liegt, ist also ebenfalls $\approx\ 2\ \Phi(c) - 1$.

Das Intervall $\left[\frac{k}{n} - \frac{c\,\sigma}{n}\ ;\ \frac{k}{n} + \frac{c\,\sigma}{n}\right]$
heißt **Vertrauensintervall für** p **zur Wahrscheinlichkeit** $\gamma = 2\ \Phi(c) - 1$.

Schätzung von p **aus dem Urnenmodell mit schwarzen und weißen Kugeln** :

Sei $\gamma = 95\% = 0{,}95$.

Dann folgt mittels der Tabelle für die **Gaußsche Summenfunktion** (siehe nächste Seite) :

$$2\ \Phi(c) - 1 = 0{,}95$$

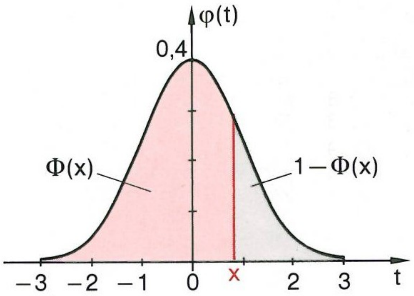
$$\Phi(c) = \frac{1 + 0{,}95}{2}$$

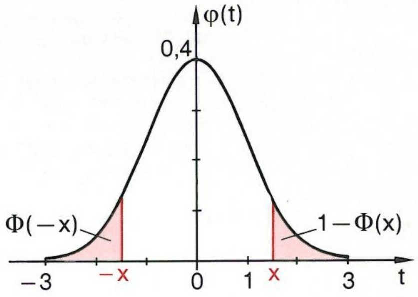
$$\Phi(c) = 0{,}975$$

$$c = 1{,}96$$

Tabelle VII: Gaußsche Summenfunktion Φ

$\Phi(x) = 0,\ldots$ $\qquad\qquad$ $\Phi(-x) = 1 - \Phi(x)$





x	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	7703	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	9996	9996	9996	9996	9996	9996	9996	9997
3,4	9997	9997	9997	9997	9997	9997	9997	9997	9997	9998

Beispiele: $\Phi(1{,}62) = 0{,}9474$, $\qquad$ $\Phi(-1{,}62) = 1 - 0{,}9474 = 0{,}0526$,
$\Phi(x) = 0{,}6772 \Rightarrow x = 0{,}46$, $\qquad$ $\Phi(x) = 0{,}3228 = 1 - 0{,}6772 \Rightarrow x = -0{,}46$

Einige besondere Werte:

x	1,2816	1,6449	1,9600	2,3263	2,5758	3,0902	3,2905
Φ(x)	0,9	0,95	0,975	0,99	0,995	0,999	0,9995

Das Experiment zeige bei $n = 100$ Ziehungen mit Zurücklegen $k = 29$ schwarze Kugeln , also $\frac{k}{n} = 0{,}29$.

$$\left|p - \frac{k}{n}\right| \le \frac{c\sigma}{n}$$

$$\left|p - \frac{k}{n}\right| \le \frac{c\sqrt{np(1-p)}}{n}$$

$$\left|p - \frac{k}{n}\right| \le c\sqrt{\frac{np(1-p)}{n^2}}$$

$$\left|p - \frac{k}{n}\right| \le c\sqrt{\frac{p(1-p)}{n}}$$

$$|p - 0{,}29| \le 1{,}96 \sqrt{\frac{p(1-p)}{100}}$$

$$(p - 0{,}29)^2 \le 1{,}96^2 \frac{p(1-p)}{100}$$

$$100 \ (p - 0{,}29)^2 \le 1{,}96^2 \ p(1-p)$$

$$100 \ (p^2 - 0{,}58\,p + 0{,}29^2) \le 1{,}96^2 \ (p-p^2)$$

$$100 \ p^2 - 58p + 100 \cdot 0{,}29^2 \le 1{,}96^2 \ p - 1{,}96^2 \ p^2$$

$$(100+1{,}96^2) \ p^2 - (58+1{,}96^2)p + 100 \cdot 0{,}29^2 \le 0$$

$$103{,}8416 \ p^2 - 61{,}8416 \ p + 8{,}41 \le 0$$

$$p_{1/2} = \frac{61{,}8416 \pm \sqrt{(-61{,}8416)^2 - 4 \cdot 103{,}8416 \cdot 8{,}41}}{2 \cdot 103{,}8416}$$

$$p_1 = 0{,}21$$

$$p_2 = 0{,}39$$

Also folgt mit $\gamma = 95\%$ Sicherheit, dass $0{,}21 \le p \le 0{,}39$ ist .

Wichtige Daten im Zusammenhang mit der Sicherheitswahrscheinlichkeit $\gamma = 2\,\Phi(c) - 1$ und dem Wert c :

$\gamma = 90\%$ $\Rightarrow$ $c = 1{,}64$

$\gamma = 95\%$ $\Rightarrow$ $c = 1{,}96$

$\gamma = 99\%$ $\Rightarrow$ $c = 2{,}58$

$\gamma = 99{,}9\%$ $\Rightarrow$ $c = 3{,}29$