KATA LOGO Mathematik – Statistik – Wahrscheinlichkeitsverteilungen - Beispiele

Verteilungen	Problemstellung	Ergebnisraum Ω	Stichprobe (n aus N)	mehrfaches Auswählen = wiederholen	Formel für P	Erwartungswert E(X)	Varianz	P
Verteilungen diskreter ZV's						$\Sigma x_i \cdot p_i$	$\Sigma (x_i - \mu)^2 \cdot p_i = E(X-\mu)^2$	
Hyper-geometrische	Stichproben ohne Zurücklegen. Dreiparametrige Verteilung. Bei Stichprobe (etwa n/N < 0,05) ähnliche Wahrscheinlichkeiten wie Binomialverteilung (B. wird dann angewendet, da einfacher). Wenn n · p · (1-p) ≥ 9 und n/N ≤ 0,05, dann Normalverteilung. Spezielle Wahrscheinlichkeitsverteilung. Da komplex, Variablen k Treffer, K Richtige und n Aussagen hinschreiben.							
	Urne mit N Objekten. k Objekte mit bestimmter Eigenschaft.	dichotom	n Objekte	nein	$\binom{K}{k} \cdot \binom{N-K}{n-k} / \binom{N}{n}$	$n \cdot \frac{M}{N}$	$n \cdot \frac{M}{N} \frac{N-M}{N} \frac{N-n}{N-1}$	p(Y)
	Urne 10 N Kugeln, 6 k rote, 4 weiße. Entnahme von 4 Kugeln (n=4). P (k = 2) ?	(r,w)	4	nein	$\binom{6}{2} \cdot \binom{10-6}{4-2} / \binom{10}{4}$			0,428
	P beim Lotto, genau k Richtige zu haben mit (k = 0,...,6)		6	nein	$\binom{6}{0} \cdot \binom{43}{6} / \binom{49}{6}$ plus ...			
	Behälter 45 N Kugeln, 20 k gelbe. P (k = 4) aus 10er Stichprobe?	(g,x,...,y)	10	nein	$\sum_k^n k \binom{20}{4} \cdot \binom{25}{6} / \binom{45}{10}$	$n \cdot \frac{M}{N} = 10 \cdot \frac{20}{45} = 0{,}44$	1,9641	0,269
	Anwendung bei Qualitätskontrollen			nein				

Letzte Aktualisierung 10.03.2015

© ANA LOGO 2015

Binomiale	Gleiche Treffer-WS bei n-maligen unabhängigen (mit Zurücklegen) Einzelversuchen (Bernoulli). Zweiparametrige Verteilung (A und A quer). 2. Frage: Für kleine Stichproben. Faustregel: n/N ≤ 0,05. Setzung: p=M/N. Voraussetzung: p = P(A) ist bekannt. Gezogene Elemente ändern P beim nächsten Zug nicht viel. Ähnliche Wahrscheinlichkeiten wie Hypergeometrische Verteilung. Wird dann angewendet, da einfacher. Verteilungsfunktion ist Σ der P von null bis zum zu errechnenden Wert: Berechnung von Hand aufwändig. Für große n erhält man über die Standardisierung der Binomialverteilung (setze z:= (X-μ)/σ) die Näherungsformel für die n → ∞ Gaußverteilung. Für Würfelergebnisse.							
	P (X=k) für P, dass unter n (unabhängigen) Versuchen genau k mal Ereignis A auftritt → k = # Treffer. p = W. einen Treffer zu ziehen. n = # Versuche.			ja: unabhängig	$\binom{n}{k} \cdot p^{k} \cdot (1-p)^{n-k}$	n · p	n · p · (1-p) Wenn n · p · (1-p) ≥ 9, dann Normalverteilung	
	Jeder 5. Autofahrer fährt zu schnell. Wie ist P für 6 kontrolliert Autos mit P(0)=0,26, P(1)=0,39, P(2)=0,082, P(4)=0,02, ...		6	ja: unabhängig	$\binom{6}{0} \cdot \binom{1}{5}$ ^ 0 · $\binom{4}{5}$ ^6	0 · 0,26 + 1 · 0,39 + 2 · 0,25 + 4 · 0,02 + ...		
	Bei Fertigung 5% (p=0.05) der Produkte fehlerhaft. Quali-Prüfung n=5 entnehmen. P für Vorfinden von 2 (k=2) defekten.		5	ja: unabhängig	$\binom{5}{1} \cdot 0{,}05^{1} \cdot (1-0{,}05)^{5-1}$			0,2036
	Bei Fertigung 5% (p=0.05) der Produkte fehlerhaft. Quali-Prüfung n=5 entnehmen. P für Vorfinden von 1 (k=1) defekten.		5	ja: unabhängig	$\binom{5}{2} \cdot 0{,}05^{2} \cdot (1-0{,}05)^{5-2}$			0,0214
	Behälter 80 Kugeln, 16 gelbe. n=5 entnehmen. P für k=3 sind gelb.		5	ja: unabhängig	$\binom{5}{3} \cdot \binom{16}{80}^{3} \cdot \binom{4}{5}^{2}$			0,0512
	253 Personen im Raum. P für k Personen, am selben Tag Geburtstag zu haben.			ja: unabhängig				
Poisson	3. Frage: n ≥ 100 und p ≤ 0,05? Setzung: p=n·p. Näherungsformel für Verteilung der seltenen Ereignisse. Spezielle Wahrscheinlichkeitsverteilung. Parametrische Klasse von Verteilungen. Bei großen Stichproben lässt sich die Binomialverteilung gut durch die Poisson-Verteilung approximieren.							

© ANA LOGO 2015

	b(n;p;k) = P(X=k) ≈ ($\mu^k \cdot e^{-\mu}$) / k!				P(X=k) = (λ^k/k!) · $e^{-\lambda}$ mit k = 0,1,...	λ	λ	
	P beim Skat: 1 von 3 Spielern hat alle 4 Buben = 0,0175. Wie P(X), dass Ereignis bei 100 Spielen ≤ 1 eintritt? μ=100·0,0175 = 1,75.				1-P(X=0) = 1- ($1,75^0 \cdot e^{-1,75}$ / 0!)	Da μ = 1,75 in Tafeln zur Poisson-verteilung nicht aufgeführt: Interpolieren.		0,8262
Diskrete Gleich-verteilung	Für Zufallsexperimente mit gleichhäufigen Ergebnissen.							
	i = 1,...,n. Es gilt { 0, falls x∈]-∞,a[und x-a/b-a, falls x∈(a,b] und 1, falls x∈]b,∞[}				P(X=xi) = 1/n mit I=1,2,...,n	1 / n · $\sum_{i=1}^{n} xi$	$\sqrt{(b-a)/\sqrt{12}}$	
Bernoulli (Null-Eins-Verteilung)	Ist dasselbe wie Binomialverteilung.							
Negative Binomial (Pascal)								

© ANA LOGO 2015

Geometrische								
					p (1-p) $^{x-1}$ mit x = 1,2,3 ...			
Logarithmische								
Verteilungen stetiger ZV's	Dichtefunktion ist erste Ableitung der monoton steigenden Verteilungsfunktion und/oder Stammfunktion F(x). f(x): ∫ -∞/+∞ f(t)dt = 1. $F_X(x)$ = P (X<x).				$\int_{-\infty}^{x} \boldsymbol{f(t)dt}$	$\int_{-\infty}^{+\infty} x$ · (fx)dx	$\int_{-\infty}^{\infty} t^2$ (f·t)dt-E(Z)²	
Stetige Gleich-verteilung	Rechteckverteilung, Uniformverteilung. Über einem endlichen Intervall [a,b], im einfachsten Fall [0,1]. Gutes Modell also, wenn Z nur Werte in [a; b] annehmen kann und mit gleicher P in alle gleich großen Teilbereiche fällt. Hat konstante Wahrscheinlichkeitsdichte auf einem Intervall (a,b): Das bedeutet: Alle Teilintervalle gleicher Länge besitzen dieselbe Wahrscheinlichkeit.							
	Verteilungsfunktion = **Stammfunktion F(x)**				**P(X<x) = 0 x < a** **P(X<x) = x-a/b-a a ≤ x < b** **P(X<x) = 1 x > b**	(a+b) / 2. E(X) = Median.	$\sqrt{(b-a)^2/12}$	
	Wahrscheinlichkeitsfunktion = **Dichtefunktion. "NV zu µ und σ".**							
	Shuttlebus fährt alle 20 min. Ein Fahrgast (der nicht weiß , zu welchen Uhrzeiten der Bus fährt) kommt zur Haltestelle. Die ZV für seine Wartezeit ist auf [0; 20]-gleichverteilt.				1 / n	(a+b) / 2	$\sqrt{(b-a)/\sqrt{12}}$	

© ANA LOGO 2015

Normal-verteilung (Gauß-Verteilung)	4. Frage: Wenn bei Poissonverteilung $\mu \geq 100$ und $p \leq 0{,}05$, also negative Zahlen fast nicht vorkommen und Verteilung symmetrisch ist, dann Normalverteilung. Wird eindeutig festgelegt durch: μ Mittelwert in der GG und σ^2 Varianz in der GG: $N(\mu,\sigma^2)$. Die meisten P-Verteilungen lassen sich bei großer Stichprobe zur NV überleiten. Gut für Prozesse in Naturwissenschaft und Einkommensverteilung, die in mehreren Faktoren unabhängig voneinander in verschiedene Richtungen wirken.							
	Verteilungsfunktion = Stammfunktion F(x)				$P(X<x) = 1 / \sigma\sqrt{2\pi} \cdot$ $\int_{-\infty}^{x} e^{(-[(t-\mu)^2]/2\sigma^2)}\, dt$ mit $\Phi(-x) = 1 - \Phi(x)$	**μ**	**σ^2**	
	Wahrscheinlichkeitsfunktion = Dichtefunktion. "NV zu μ und σ"				**Dichtefunktion f(x):** $[1 / 2\pi\sigma] \cdot e^{-[(x-\mu^2)/2\sigma^2]}$	$P(\mu-\sigma \leq Z \leq \mu+\sigma) = 2\,\phi(1) - 1 = 0{,}6827$ $P(\mu-2 \leq Z \leq \mu+2\sigma) = 2\,\phi(2) - 1 = 0{,}9545$ $P(\mu-3\sigma \leq Z \leq \mu+3\sigma) = 2\,\phi(3) - 1 = 0{,}9973$	siehe nebenan	
Standardisierte Normal-verteilung								
	Verteilungsfunktion groß ϕ(a) = gaußsche Summenfunktion				$P(X<x) = 1 / \sigma\sqrt{2\pi} \cdot$ $\int_{-\infty}^{x} e^{(-[(t-\mu)^2]/2\sigma^2)}\, dt$ mit $\Phi(-x) = 1 - \Phi(x)$	**0**	**1**	
	Wahrscheinlichkeitsfunktion klein ϕ(a) = Dichtefunktion. Hat keine elementare Stammfunktion.				$[1 / \sqrt{2\pi\sigma}] \cdot e^{-x^2/2}$			
	Nach Umformung (s.o.) von X gemäß $U = (X-\mu)/\sigma$, erhält man standardisierte NV: ϕ (u) mit Verteilungsfunktion Φ (u).				$[1 / \sqrt{2\pi}] \cdot e^{(-0{,}5 \cdot u^2)}$ mit $\Phi(-u) = 1 - \Phi(u)$			
	Nach Umformung (s.o.) von X gemäß $U = (X-\mu)/\sigma$, erhält man standardisierte NV: ϕ (u) mit Verteilungsfunktion Φ (u):				$[1 / \sqrt{2\pi}] \cdot$ $\int_{-\infty}^{u} e^{(-0{,}5 \cdot t^2)}\, dt$ mit $\Phi(-u) = 1 - \Phi(u)$			

© ANA LOGO 2015

Exponential e^{λ}	Verteilungsfunktion ist die Stammfunktion F(x). Durch Ableitung f(x) kommt man zu ihrer Dichtefunktion. Einparametrige Verteilung: Wird durch Parameter λ bestimmt mit Werten 0 (asymptotischer Graph) und unendlich (exponentialer Graph). Modell für Wartezeit auf ein bestimmtes Ereignis, wenn Eintreten des Ereignisses nicht wahrscheinlicher oder unwahrscheinlicher wird, wenn es eine Zeitlang nicht eingetreten ist. Obige Bedingung besagt dann nämlich: P, dass Ereignis in einem Zeitraum [0; t] eintritt, ist genauso groß wie die (bedingte) P dafür, dass es im Zeitraum [t0; t0+t] eintritt, wenn bekannt ist, dass es zum Zeitpunkt t0 noch nicht eingetreten ist. Über einem halbseitig unendlichen Intervall, üblicherweise als [0,∞] angenommen.							
	Verteilungsfunktion und Stammfunktion F(x) kumuliert				$\int_0^x f\lambda(t)dt$ **= 1-e$^{-\lambda \cdot x}$** **für x≥0, und 0 für x<0}**	μ = 1/λ Median = ln(2) / λ	Varianz 1/λ² und STA: 1/λ	
	Dichtefunktion $f_{\lambda}(x) = \lambda e^{-\lambda x}$ wenn x ≥ 0 und 0 wenn x <0. Bei Zeitthemen auch t genannt.				**$f_{\lambda}(x) = \lambda e^{-\lambda x}$** **wenn x ≥ 0 und 0 wenn x <0**			
	P für Brenndauer Glühbirne 1.000-2.000h bei λ= 0,001: Verteilungsfunktion: 1-e$^{\lambda \cdot 0,001 \cdot x}$				P (1.000 ≤ Z ≤ 2.000) = F(2.000) - F(1.000) = $(1-e^{-0,001 \cdot 2.000}) - (1-e^{-0,001 \cdot 1.000}) = e^{-1} - e^{-2}$			
Chi²	Wenn ZV = (Σ voneinander unabhängiger quadrierter Standardnormalvariablen (siehe Standardnormalverteilung)), dann Chi². Asymetrisch. Verwendung beim Chi²-Anpassungstest, Chi²-Homogenitätstest und Chi²-Unabhängigkeitstest. Stichprobenverteilung v.a für die Schätzung von Verteilungsparametern, z.B. der Varianz. Über einem halbseitig unendlichen Intervall, üblicherweise als [0,∞] angenommen.							
Student t-Verteilung	Fast so wie Normalverteilung, aber schmäler.							

© ANA LOGO 2015

F-Verteilung (Fisher-Verteilung)	Über einem halbseitig unendlichen Intervall, üblicherweise als [0,∞] angenommen:							
Logarithmische Normal	Auch Log-Normalverteilung. Über einem halbseitig unendlichen Intervall, üblicherweise als [0,∞] wie bei Verteilung von Zeiten angenommen. Viele Verteilungen in der Natur laufen als schiefe, linkssteile Verteilung rechts flach aus. Erklärung: Merkmal kann oft bestimmten Schrankenwert nicht über- bzw. unterschreiten und kann nach dieser Seite hin in seiner Variationsmöglichkeit gehemmt sein. Durch Logarithmieren Überführung des Berecihes [0,1] in [-∞,0]. Linker Teil wird stark gestreckt, und rechter stark gestaucht, v.a. bei großer STA im Vergleich zu x quer. Viele Zufallsgrößen wirken multiplikativ zusammen. Biologie: Empfindlichkeit von Tieren einer Art gegenüber Pharmaka.							
						$e^{\mu + (\sigma^2 / 2)}$. Median e^{μ}	$e^{2\mu + \sigma^2} = e^{\sigma^2} - 1$	
Gamma	Über einem halbseitig unendlichen Intervall, üblicherweise als [0,∞] angenommen:							
Beta	Über einem endlichen Intervall [a,b], im einfachsten Fall [0,1].							

Letzte Aktualisierung 10.03.2015

© ANA LOGO 2015

Logistische								
Weibull	Über einem halbseitig unendlichen Intervall, üblicherweise als [0,∞] angenommen.							
Pareto	Über einem halbseitig unendlichen Intervall, üblicherweise als [0,∞] angenommen.							
Benford								

© ANA LOGO 2015