Stochastik für die Naturwissenschaften

Dr. C.J. Luchsinger

6. Ausgewählte Verteilungen (Distributions)

Dies ist Ihr wichtigstes Kapitel für die Prüfung! Alle Resultate zu allen Verteilungen sind hier zusammengefasst und dürfen benutzt werden, auch wenn wir sie nicht in der Vorlesung benutzt haben (zB $E[\chi_n^2] = n$, siehe später). Alle ausgedruckten Tabellen der Wahrscheinlichkeits- und Verteilungsfunktionen zu Binomial, $\mathcal{N}$, χ_n^2, $F_{m,n}$ und t_n hier einordnen!

Wir präsentieren hier viele Resultate, welche wir aus Zeitgründen nicht alle nachprüfen. Viele der nachfolgenden Resultate sind bereits bekannt und werden hier nur kompakt zusammengefasst, sodass wir relativ rasch voranschreiten werden; bitte zu Hause in Ruhe durchlesen. Es kommen auch Verteilungen vor, welchen wir bis jetzt noch nicht begegnet sind. Wir werden diese Teile in unterer Zusammenfassung nochmals repetieren, sobald wir auf diese Verteilungen stossen; es ist also nicht gravierend, wenn man einzelne Teile (noch) nicht versteht.

Die StudentInnen sollten von Kapitel 6 wissen, wie Wahrscheinlichkeitsfunktionen oder Dichten der wichtigsten Verteilungen aussehen. Wir werden dies in der Statistik (Ablehnungsbereiche bei Tests) intensiv brauchen.

6.1 Diskrete Verteilungen

6.1.1 Bernoulli Be(p); R: binom mit $n = 1$

X kann 2 Werte annehmen: 0 und 1 (alternativ auch -1 und $+1$). $P[X = 1] = p$ (Erfolg) und $P[X = 0] = 1 - p$ (Misserfolg). $E[X] = p$ und $V[X] = p(1 - p)$. Mit $0 \le k \le 1$ haben wir

$$P[X = k] = p^k(1 - p)^{1-k}.$$

Dies ist in der Tat gleichwertig wie oben (rechne), zielt aber bereits auf die Bin(n,p).

6.1.2 Binomial Bin(n,p); Storrer Kapitel 38, 51.1; R: binom

Seien X_i, $1 \leq i \leq n$, n iid Be(p)-Zufallsgrössen. Sei $Y := \sum_{i=1}^{n} X_i$. Dann hat Y per Definitionem die Binomialverteilung mit Parametern n und p; Bin(n,p). $E[Y] = np$ (Bemerkung zu Lemma 5.3) und $V[Y] = np(1-p)$ (Bemerkung zu Lemma 5.5). $0 \leq k \leq n$:

$$P[Y = k] = \binom{n}{k} p^k (1-p)^{n-k}.$$

Einsatz: Anzahl Erfolge (k) bei n unabhängigen Versuchen mit Erfolgswahrscheinlichkeit p.

6.1.3 Geometrisch Ge(p); Storrer (37.12 b)); R: geom

Bei dieser Verteilung interessieren wir uns für den Zeitpunkt des ersten Erfolgs. Z ist eine Zufallsgrösse auf den natürlichen Zahlen ohne die Null. Wir haben

$$P[Z = k] = p(1-p)^{k-1}.$$

Es gilt: $E[Z] = 1/p$ und $V[Z] = (1-p)/p^2$. Die Ge(p) hat die Eigenschaft, dass sie die einzige diskrete Zufallsgrösse ist, welche *gedächtnislos* ist: mit $n > m > 0$ gilt hier

$$P[Z > n | Z > m] = P[Z > (n-m)].$$

"Gegeben, es hat schon $m = 1000$ Würfe ohne 6 (=Erfolg) gegeben, wie gross ist die Wahrscheinlichkeit, dass es sogar insgesamt mindestens $n = 1004$ Würfe ohne 6 geben wird? Also dies ist nur noch von der Differenz $n - m = 4$ abhängig. Wie lange es bereits keine 6 gegeben hat, ist egal!"

Es sei noch erwähnt, dass die geometrische Verteilung in gewissen Lehrbüchern so definiert wird, dass man die Anzahl Misserfolge bis zum ersten Erfolg zählt. Dann nimmt die Ge(p) Werte auf den natürlichen Zahlen *inklusive* die 0 an. Die Resultate sind analog aber leicht komplizierter.

Woher kommt der Name "Geometrisch"?

6.1.4 Poisson Po(λ); Storrer Kapitel 39, 51.2; R: pois

Eine Zufallsgrösse X ist poissonsch, wenn Sie Werte auf den natürlichen Zahlen inklusive 0 annimmt und zwar mit folgenden Wahrscheinlichkeiten:

$$P[X = k] = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Es gilt: $E[X] = V[X] = \lambda$ (E: Beispiel 3 in 5.1). Die Motivation für die Poissonverteilung folgt in Kapitel 7.

6.2 Stetige Verteilungen

6.2.1 Uniform $U[a, b]$; R: unif

Die einfachste stetige Verteilung ist die Uniform-Verteilung: Eine Zufallsgrösse U ist per Definitionem auf dem Intervall $[a, b]$ uniform verteilt, wenn U folgende Dichtefunktion hat:

$$f(u) = (b - a)^{-1},$$

wobei dann natürlich $a \leq u \leq b$ zu gelten hat. Ausserhalb von $[a, b]$ ist die Dichte gleich null. Es gilt $E[U] = (a + b)/2$ (2. Beispiel von 5.1) und $V[U] = (b - a)^2/12$ (Beispiel 6 in 5.1):

6.2.2 (Negativ-) Exponential Exp(λ); Storrer (40.8); R: exp

Eine Zufallsgrösse X mit Dichtefunktion

$$f(x) = \lambda e^{-\lambda x}, \;\; x \geq 0,$$

heisst exponentialverteilt mit Parameter (oder Rate) λ; Exp(λ). Es gilt $E[X] = 1/\lambda$ und $V[X] = 1/\lambda^2$. Es macht Sinn, dass die erwartete Zeit bis zum (nächsten) Ereignis umgekehrt proportional zur Rate sein muss. Modell für: Zeit bis radioaktiver Zerfall, "wann geht eine Glühbirne kaputt?", Zwischenzeit bei der Ankunft von KundInnen in einem Geschäft und vieles mehr.

Die Exp(λ) hat die Eigenschaft, dass sie die einzige stetige Zufallsgrösse ist, welche *gedächtnislos* ist: mit $t > s > 0$ gilt hier

$$P[X > t | X > s] = P[X > (t - s)].$$

"Gegeben, es hat schon $s = 1000$ Sekunden keinen Atomzerfall gegeben, wie gross ist die Wahrscheinlichkeit, dass es sogar insgesamt mindestens $t = 1004$ Sekunden keinen Atomzerfall geben wird? Also dies ist nur noch von der Differenz $t - s = 4$ Sekunden abhängig. Wie lange es bereits keinen Atomzerfall gegeben hat, ist egal!"

Gemeinsamkeiten von Ge(p) und Exp(λ) und Konsequenzen (Signalwort ist (Übergangs- oder Warte-)Zeit bis):

6.2.3 Normal $\mathcal{N}(\mu, \sigma^2)$; auch Gauss-, Glocken-, Bell-, Forrest Gump-Verteilung, Storrer Kapitel 41; neu zu lesen: (41.3)-(41.7), 51.3; R: norm

Vorsicht: Storrer $\mathcal{N}(\mu, \sigma)$, Luchs $\mathcal{N}(\mu, \sigma^2)$

Wegen des zentralen Grenzwertsatzes (siehe Kapitel 7) ist die Normalverteilung sehr wichtig: Mit Dichtefunktion

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \; x \in \mathbb{R},$$

gilt $E[X] = \mu$ (Beispiel 4 in 5.1) und $V[X] = \sigma^2$ (Beispiel 8 in 5.1).

"Z-Transform" [Beweis Storrer (41.3)]: Wenn X eine $\mathcal{N}(\mu, \sigma^2)$-Verteilung hat, dann hat

$$Z := \frac{X - \mu}{\sigma} \tag{Z – Transform}$$

eine $\mathcal{N}(0, 1)$-Verteilung (welche in Büchern und Libraries von Statistik-Paketen abgelegt ist). $\mathcal{N}(0, 1)$ nennen wir "Standard-Normalverteilung". 4 Bemerkungen:

Wir machen dazu noch 3 Bilder (Zentrieren und Normalisieren = Standardisieren):

Sei X eine $\mathcal{N}(10, 4)$-Zufallsgrösse. Formen Sie $P[X > 13]$ soweit um, dass Sie den Wert in einer Tabelle ablesen können.

Storrer Aufgabe 41-3 und 41-6.

Approximative Formeln aus der Praxis: sei X eine $\mathcal{N}(0,1)$-Zufallsgrösse. Dann gelten $P[|X| > 1] \doteq 1/3$ und $P[|X| > 2] \doteq 5\%$, zwei Bilder:

Allgemein gilt wegen der Z-Transform (rechnen Sie es selber nach!): Eine Normalverteilung hat etwa 2/3 ihrer Wahrscheinlichkeit innerhalb 1 Standardabweichung vom Mittelwert entfernt (genauer, siehe Tabelle: %) und sogar 95 % ihrer Wahrscheinlichkeit innerhalb von 2 Standardabweichungen vom Mittelwert entfernt:

6.2.4 χ^2; Storrer Kapitel 47, 51.5; R: chisq

Diese Verteilung ist in der Statistik zentral wichtig und verdankt ihre Bedeutung weitgehend dem zentralen Grenzwertsatz (Kapitel 7), insbesondere, dass man in Modellen der Datenanalyse Fehlerterme häufig normalverteilt modelliert. Wenn $(X_i)_{i=1}^n$ iid $\mathcal{N}(0,1)$-verteilt sind, dann ist

$$\sum_{i=1}^n X_i^2$$

χ_n^2-verteilt (sprich "Chiquadrat mit n Freiheitsgraden" (degree of freedom, df)):

Wenn $(Y_i)_{i=1}^n$ iid $\mathcal{N}(\mu, \sigma^2)$-verteilt sind, dann hat wegen der Z-Transformation

$$\sum_{i=1}^n \frac{(Y_i - \mu)^2}{\sigma^2}$$

die χ_n^2-Verteilung. Ein wenig überraschend: Mit

$$S^2 := \sum_{i=1}^n (Y_i - \overline{Y})^2,$$

wo $\overline{Y} := \sum_{i=1}^n Y_i/n$, hat S^2/σ^2 die χ_{n-1}^2-Verteilung. Eine Bauernregel sagt: ein df geht verloren pro Parameter, den man schätzt (μ wird mit $\overline{Y}$ geschätzt). Die Dichte der χ_n^2-Verteilung ist

$$f(x) := \frac{x^{n/2-1}e^{-x/2}}{\Gamma(n/2)2^{n/2}},\ x \geq 0. \qquad \text{(Monster I)}$$

$E[\chi_n^2] = n; V[\chi_n^2] = 2n$. Bilder zur Dichte in Storrer (47.2).

In Anbetracht der Definition ist es nicht überraschend, dass, wenn X eine χ_a^2-Zufallsgrösse und Y eine χ_b^2-Zufallsgrösse ist und X und Y unabhängig voneinander sind, dann ist $W := X + Y$ eine χ_{a+b}^2-Zufallsgrösse.

6.2.5 F; R: f

Seien U und V zwei unabhängige, χ_m^2-, bzw. χ_n^2-verteilte Zufallsgrössen. Dann ist der Ausdruck

$$W := \frac{U/m}{V/n}$$

F-verteilt mit Parametern m, n: $F_{m,n}$. Diese Zufallsgrössen kommen in Kapitel 10 vor. Für $x \geq 0$ haben wir die Dichtefunktion (Monster II)

$$f(x) = \frac{\Gamma[(m+n)/2](m/n)^{m/2}}{\Gamma(m/2)\Gamma(n/2)} \frac{x^{m/2-1}}{(1 + mx/n)^{(m+n)/2}}$$

Es gilt $E[W] = n/(n-2)$ (falls $n > 2$),

$V[W] = [2n^2(m+n-2)]/[m(n-2)^2(n-4)]$.

6.2.6 t; Storrer Kapitel 45-46, 51.4; R: t

Sei Y eine $\mathcal{N}(0,1)$-Zufallsgrösse und Z eine χ^2_n-Zufallsgrösse; Y unabhängig von Z.

$$T_n := \frac{Y}{\sqrt{Z/n}}$$

ist dann fast $\mathcal{N}(0,1)$-verteilt, aber nicht genau, die genaue Verteilung ist Student-t-Verteilung mit n df. Die Dichte der t_n-Verteilung ist (Monster III)

$$f(x) = \frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\Gamma(n/2)} \frac{1}{(1+x^2/n)^{(n+1)/2}},\ x \in \mathbb{R}.$$

Es gilt $E[t_n] = 0$ (falls $n > 1$ und falls $n = 1$ existiert der Erwartungswert nicht), $V[t_n] = n/(n-2)$, sobald $n > 2$. Man sieht sofort, dass "$\sqrt{F_{1,n}} = |t_n|$". $t_\infty = \mathcal{N}(0,1)$, siehe Tabellen in (51.4).

3 Arten von Tabellen im Storrer, bzw im Web (F-Tabelle):

$\mathcal{N}$:

t und chi-Quadrat:

F:

6.3 Zusammenhänge bei Verteilungen I: $\sum$ von *unabhängigen* Zufallsgrössen

Die Summanden in den folgenden Summen müssen immer unabhängig sein. Wenn nicht anders erwähnt, fordern wir auch die gleiche Verteilung. Mit kompakten Formeln wie

$$\sum_1^n Be(p) = Bin(n,p)$$

weiter unten ist eigentlich gemeint: Seien $X_1, \ldots, X_n$ iid Be(p)-verteilt. Dann hat

$$Y := \sum_{i=1}^n X_i$$

eine Bin(n,p)-Verteilung.

* Summe von Bernoulli ist Binomial

$$\sum_1^n Be(p) = Bin(n,p)$$

* Grosser Bruder hiervon: Summe von Binomial ist Binomial (p muss immer gleich sein!)

$$\sum_1^n Bin(n_i,p) = Bin(\sum_1^n n_i, p)$$

* Summe von Normal ist Normal (Summanden müssen nicht mal identisch verteilt sein)

$$\sum_{i=1}^n \mathcal{N}(\mu_i, \sigma_i^2) = \mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$$