Einführung in die Statistik

Dr. C.J. Luchsinger

4 Ausgewählte Verteilungen

* diskret: Bernoulli, Binomial, Geometrisch, Negativ-Binomial, Poisson

* stetig: Uniform, (Negativ-)Exponential, Gamma, Normal, Cauchy, χ^2, F, t

* Wertebereich, Verteilung (Wahrscheinlichkeitsfunktionen / Dichten), E, V, R-Befehl, Einsatz

Es gibt bekanntlich ∞ viele Verteilungen. Also gäbe es auch ∞ viele Erwartungswerte und Varianzen etc. zu berechnen. Wir präsentieren hier viele Resultate, welche wir aus Zeitgründen nicht alle nachprüfen. Man beachte, dass sowohl bei diskreten wie auch bei stetigen Zufallsgrössen die Wahrscheinlichkeitsfunktionen resp. Dichten fast immer nach dem gleichen Schema aufgebaut sind:

Normierungskonstante mal informativer Teil, zum Beispiel

$$\frac{1}{\sqrt{2\pi}\sigma} * e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

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 4 wissen, wie Wahrscheinlichkeitsfunktionen oder Dichten der wichtigsten Verteilungen aussehen. Wir werden dies in Kapitel 6 (Ablehnungsbereiche) intensiv brauchen.

4.1 Bemerkung zu den R-Befehlen

Wir werden immer nur den ”Stamm” des Befehls in R angeben. Davor kann man dann ”d” (z.B. dnorm(0) für Dichtefunktion beim Punkt 0 einer $\mathcal{N}(0,1)$-Zufallsgrösse) setzen. ”p” gibt die Verteilungsfunktion, ”q” die Quantilfunktion (”bis wo muss ich gehen, dass 40 % der Verteilung hinter mir liegen?”) und ”r” gibt dann eine Stichprobe. Lesen Sie dazu auch Kapitel 2 im Dalgaard und führen Sie dazu die Befehle in R aus (im Lam ist es Sektion 9.1.1).

$> dnorm(0)$

[1]0.3989423

$> pnorm(0)$

[1]0.5

$> qnorm(0.5)$

[1]0 (vgl. ein Schritt weiter oben)

$> qnorm(0.95)$

[1]1.644854 (das berühmte ”1.64” aus der Statistik)

$> qnorm(0.975)$

[1]1.959964 (das berühmte ”1.96” aus der Statistik)

4.2 Diskrete Verteilungen

4.2.1 Bernoulli Be(p)

X kann 2 Werte annehmen: 0 und 1 (alternativ auch -1 und $+1$, siehe z.B. Aufgabe auf Blatt 7). $P[X = 1] = p$ (Erfolg) und $P[X = 0] = 1 - p$ (Misserfolg). $E[X] = p$ und $V[X] = p(1 - p)$. **R-Befehl:** binom mit $n = 1$. Mit $0 \leq k \leq 1$ haben wir

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

4.2.2 Binomial Bin(n,p); Krengel § 2: 2.4

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). $E[Y] = np$ (Aufgabe auf Blatt 6) und $V[Y] = np(1 - p)$ (Bemerkung zu Lemma 3.8), **R-Befehl:** binom. $0 \leq k \leq n$:

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

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

4.2.3 Geometrisch Ge(p); Krengel § 2: 2.4

Seien wieder X_i, $i \geq 1$, iid Be(p)-Zufallsgrössen. Wir interessieren uns für den Zeitpunkt des ersten Erfolgs: $Z := \min\{i | X_i = 1\}$. Z ist eine Zufallsgrösse auf den natürlichen Zahlen ohne die Null. Z hat per Definitionem eine geometrische Verteilung Ge(p) mit Parameter p. Es gilt: $E[Z] = 1/p$ und $V[Z] = (1 - p)/p^2$, **R-Befehl:** geom.

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

Die Ge(p) hat die für die angewandte Stochastik zentral wichtige Eigenschaft, dass sie die einzige diskrete Zufallsgrösse ist, welche *gedächtnislos* ist: mit $n > m > 0$ gilt hier (vgl. Aufgabenblatt 9)

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

"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 - das selbe gilt auch für eine alternative Definition der NB(n,p) nachfolgend.

4.2.4 Negativ Binomial NB(n,p); Krengel § 2: 2.4

Seien Z_i, $1 \leq i \leq n$, n iid Ge(p)-Zufallsgrössen. Sei $W := \sum_{i=1}^n Z_i$. Dann hat W per Definitionem die Negativ-Binomialverteilung mit Parametern n und p; NB(n,p). Die NB(n,p)-Verteilung beschreibt die Zeit, bis es n Erfolge gegeben hat. $E[W] = n/p$ (folgt aus Linearität des Erwartungswerts und 4.2.3) und $V[W] = n(1-p)/p^2$ (folgt aus Lemma 3.8 b) und 4.2.3), **R-Befehl:** nbinom. $w \geq n$

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

4.2.5 Poisson Po(λ); Krengel § 5: 5.4

Die Motivation für die Poissonverteilung folgt in der Vorlesung Angewandte Stochastik. Eine Zufallsgrösse V ist poissonsch, wenn Sie Werte auf den natürlichen Zahlen inklusive 0 annimmt und zwar mit folgenden Wahrscheinlichkeiten:

$$P[V = v] = e^{-\lambda} \frac{\lambda^v}{v!}.$$

Es gilt: $E[V] = V[V] = \lambda$ (Beispiel 3 in 3.1 und Bemerkung 3 zu Lemma 3.7), **R-Befehl:** pois.

4.3 Stetige Verteilungen

4.3.1 Uniform $U[a, b]$; Krengel § 10: 10.2

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 3.1) und $V[U] = (b - a)^2/12$ (Beispiel 6 in 3.1), **R-Befehl:** unif.

4.3.2 (Negativ-) Exponential Exp(λ); Krengel § 10: 10.2

Eine Zufallsgrösse X mit Dichtefunktion

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

heisst exponentialverteilt mit Parameter λ; Exp(λ). Die gleiche Wahl des Parameters wie bei der Poissonverteilung ist *nicht* zufällig! Es gilt $E[X] = 1/\lambda$ und $V[X] = 1/\lambda^2$ (Aufgabenblatt 6), **R-Befehl:** exp. Modell für: radioaktiver Zerfall, "wann geht eine Glühbirne kaputt?", Zwischenzeit bei der Ankunft von KundInnen in einem Geschäft und vieles mehr; mehr dazu in der Vorlesung "Angewandte Stochastik".

Die Exp(λ) hat die für die angewandte Stochastik zentral wichtige Eigenschaft, dass sie die einzige stetige Zufallsgrösse ist, welche *gedächtnislos* ist: mit $t > s > 0$ gilt hier (vgl. Aufgabenblatt 9)

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

"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!"

4.3.3 Gamma(n, λ); Krengel § 14: Anhang

Seien X_i, $1 \leq i \leq n$, n iid Exp(λ)-Zufallsgrössen. Sei $Y := \sum_{i=1}^n X_i$. Dann hat Y per Definitionem die Gammaverteilung mit Parametern n und λ; Gamma(n,λ). $E[Y] = n/\lambda$ (folgt aus Linearität des Erwartungswerts und 4.3.2) und $V[Y] = n/\lambda^2$ (folgt aus Lemma 3.8 b) und 4.3.2), **R-Befehl:** gamma.

$$f(y) = \frac{y^{n-1}e^{-\lambda y}\lambda^n}{\Gamma(n)},\ y \geq 0.$$

Mit $n = 1$ wird dies in der Tat zu einer Exponentialverteilung. In Anbetracht von 4.3.2 ist das ein Modell für "wann zerfällt das n-te Atom?", geht die n-te Glühbirne kaputt, kommt der n-te Kunde. In der Statistik hat man mit der Gamma-Verteilung auch ein relativ flexibles Modell, um Einkommensverteilungen zu modellieren (2 frei wählbare Parameter n, λ).

4.3.4 Normal N(μ, σ^2), auch Gauss-, Glocken-, Bell-, Forrest Gump-Verteilung; Krengel § 10: 10.2 und § 5: 5.2; Tabelle II hinten

Wegen des zentralen Grenzwertsatzes 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 3.1) und $V[X] = \sigma^2$ (Beispiel 8 in 3.1), **R-Befehl:** norm.

ACHTUNG: FÜR DIE FOLGENDEN BEIDEN RESULTATE SETZEN WIR NORMALVERTEILUNG VORAUS!

1. "Z-Transform" [Beweis auf Übungsblatt 9]: Wenn X eine $\mathcal{N}(\mu, \sigma^2)$-Verteilung hat, dann hat

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

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

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

2. Approximative Formeln für beliebige $\sigma > 0$: sei X eine $\mathcal{N}(0,\sigma^2)$-Zufallsgrösse. Dann gelten

$$P[|X| > \sigma] \doteq \frac{1}{3}$$

und

$$P[|X| > 2\sigma] \doteq 5\%.$$

Allgemein: 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:

4.3.5 Cauchy

Wir haben in Aufgabenblatt 5 die Cauchy-Verteilung in einem physikalischen Experiment erhalten. Gleichzeitig ist sie die t_1-Verteilung (siehe 4.3.8 weiter unten). Neben diesen praktischen Anwendungen ist sie für die Theorie der Statistik und Wahrscheinlichkeitsrechnung gut als Verteilung mit viel Gewicht in den Enden (Langschwänzigkeit, heavy tails) im Vergleich zur Normalverteilung. Es gilt insbesondere $E[|X|] = \infty$, wie man aus unterer Dichte sofort sieht:

$$f(x) = \frac{d}{\pi(d^2 + (x-m)^2)}, \; x \in \mathbb{R}.$$

Dabei ist m der Median und d ein Skalenparameter. Die Varianz existiert auch nicht (Grund z.B. Aufgabe in Honours-Programm auf Blatt 7). **R-Befehl:** cauchy.

4.3.6 χ^2; Krengel § 14: 14.1; Tabelle III hinten

Diese Verteilung ist in der Statistik zentral wichtig und verdankt ihre Existenz weitgehend dem zentralen Grenzwertsatz (Kapitel 5), insbesondere, dass man in Modellen der Datenanalyse Fehlerterme 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 (hier μ wird mit $\overline{Y}$ geschätzt). Wenn wir mit

3.3 vergleichen, sieht man, dass S^2/n sich offenbar als Schätzer für σ^2 aufdrängt (falls σ unbekannt); mehr dazu in Kapitel 7. Die Dichte der χ^2_n-Verteilung ist

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

$E[\chi^2_n] = n; V[\chi^2_n] = 2n$. Wenn man die χ^2_n-Verteilung mit der Gamma-Verteilung vergleicht, sieht man, dass jede zweite χ^2_n-Verteilung eine Gamma-Verteilung mit Parameter $\lambda = 0.5$ ist, wenn man n erhöht. **R-Befehl:** chisq.

4.3.7 F; Krengel § 14: 14.2; Tabelle IV hinten

Seien U und V zwei unabhängige, χ^2_m-, bzw. χ^2_n-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 der Statistik vor, wenn man testen will, ob die Varianzen von 2 unabhängigen Stichproben gleich sind oder nicht (in viel komplexeren statistischen Untersuchungen kommt $F_{m,n}$ auch (relativ unerwartet) vor).

$$f(w) = \frac{\Gamma[(m+n)/2](m/n)^{m/2}}{\Gamma(m/2)\Gamma(n/2)} \frac{w^{m/2-1}}{(1+mw/n)^{(m+n)/2}},\ w \geq 0, \qquad \text{(Monster II)}$$

ist die Dichtefunktion von W. Es gilt $E[W] = n/(n-2)$ (falls $n > 2$), $V[W] = [2n^2(m+n-2)]/[m(n-2)^2(n-4)]$. **R-Befehl:** f.

4.3.8 t; Krengel § 14: 14.1; Tabelle I hinten

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

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

Es gilt $E[t_n] = 0$ (falls $n > 1$ und falls $n = 1$ existiert der Erwartungswert nicht), $V[t_n] = n/(n-2)$ sobald $n > 2$. **R-Befehl:** t. Man sieht sofort, dass "$\sqrt{F_{1,n}} = |t_n|$". Des weiteren: sind X_1, X_2 zwei iid $\mathcal{N}(0,1)$-Zufallsgrössen, so ist X_1/X_2 ebenfalls t_1- (und damit Cauchy-) verteilt. $t_\infty = \mathcal{N}(0,1)$.

Da wir in diesem Skript keine Bilder/Graphiken haben, sollte jedeR StudentIn sich eine Stunde Zeit nehmen, und in R bei obigen Verteilungen die Wahrscheinlichkeitsfunktionen bzw. Dichten anschauen. Variieren Sie dabei E, V. Lesen Sie dazu in Dalgaard Kapitel 2. Weiter ist hier ein Besuch der "Online-Demos zu Verteilungen - ein Muss" (Link auf der Homepage der Vorlesung), also

www.fernuni-hagen.de/neuestatistik/applets/index2.htm

sehr empfohlen!

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

4.4.1 Theoretische Grundlagen; Krengel § 11: 11.3

Wir werden uns in diesem Teil mit folgender Frage auseinandersetzen: "Wenn X und Y zwei unabhängige Zufallsgrössen sind, wie ist dann $X+Y$ verteilt"?

4.4.1.1 Diskreter Fall

Wir benützen im ersten Schritt die FTW aus Kapitel 1 und danach die Unabhängigkeit von X und Y:

$$\begin{aligned}
P[X+Y=a] &= \sum_i P[X+Y=a|Y=i]P[Y=i] \\
&= \sum_i P[X+i=a|Y=i]P[Y=i] \\
&:= \sum_i \frac{P[X+i=a, Y=i]}{P[Y=i]} P[Y=i] \\
&= \sum_i \frac{P[X+i=a]P[Y=i]}{P[Y=i]} P[Y=i] \\
&= \sum_i P[X+i=a]P[Y=i] = \sum_i P[X=a-i]P[Y=i].
\end{aligned}$$

Auf Blatt 9 ist dazu eine Aufgabe zu lösen.

4.4.1.2 Stetiger Fall

Wir benutzen hier den Satz von Fubini, welcher erst im dritten Semester bereitgestellt wird (Vertauschung der Integrationsreihenfolge).

$$\begin{aligned}
P[X+Y \le a] &:= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_{X,Y}(x,y)dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a} f_{X,Y}(x-y,y)dxdy \\
&= \int_{-\infty}^{a} \int_{-\infty}^{\infty} f_{X,Y}(x-y,y)dydx \\
&= \int_{-\infty}^{a} \int_{-\infty}^{\infty} f_X(x-y)f_Y(y)dydx.
\end{aligned}$$

Also ist $\int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$ die Dichte von $X+Y$ an der Stelle a.

Als Beispiel geben wir jetzt die Summe von 2 unabhängigen normalverteilten Zufallsgrössen an, welche *nicht* die gleiche Verteilung haben müssen. Sei X $\mathcal{N}(\mu_1, \sigma_1^2)$ verteilt und Y $\mathcal{N}(\mu_2, \sigma_2^2)$ verteilt, $X \coprod Y$. Wir wollen zeigen, dass $X + Y$ eine $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ Verteilung hat (durch Induktion haben wir sofort das analoge Resultat für endliche Summen). Man mache sich bitte klar, dass wegen der Linearität des Erwartungswertes und Lemma 3.8 b) der neue Mittelwert $\mu_1 + \mu_2$ und die Varianz $\sigma_1^2 + \sigma_2^2$ sein *müssen*. Wir müssen nur noch zeigen, dass $X + Y$ normalverteilt ist.

1. Wir wollen zuerst zeigen, dass wir oBdA $\mu_1 = \mu_2 = 0$ setzen dürfen: $X' := X - \mu_1$ ist $\mathcal{N}(0, \sigma_1^2)$ verteilt und $Y' := Y - \mu_2$ ist $\mathcal{N}(0, \sigma_2^2)$ verteilt. Wenn wir gezeigt haben, dass $X' + Y'$ eine $\mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$ Verteilung besitzt, so hat $X + Y$ eine $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ Verteilung. Also können wir gleich oBdA voraussetzen, dass $\mu_1 = \mu_2 = 0$.

2. Die Dichte $h(a)$ von $X + Y$ muss also sein:
$$h(a) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{(a-y)^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2}\right]} dy.$$
Hausaufgabe: nachrechnen, dass gilt:
$$\frac{(a-y)^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} = \left[\frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1\sigma_2} y - \frac{\sigma_2}{\sigma_1\sqrt{\sigma_1^2 + \sigma_2^2}} a\right]^2 + \frac{a^2}{\sigma_1^2 + \sigma_2^2}.$$
Wir definieren jetzt:
$$z(y) := \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1\sigma_2} y - \frac{\sigma_2}{\sigma_1\sqrt{\sigma_1^2 + \sigma_2^2}} a.$$
Dann gilt:
$$\frac{dz}{dy} = \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1\sigma_2}$$
und damit
$$dy = \frac{dz\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}.$$
Die Substitution ergibt damit
$$h(a) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{1}{2}\frac{a^2}{\sigma_1^2 + \sigma_2^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz.$$

□

3. Auf Aufgabenblatt 9 ist dazu eine Simulation zu machen.

4.4.2 Summen von Binomial, Geometrisch, Poisson, Exponential, Normal

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 $Y := \sum_1^n X_i$ 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 Geometrisch ist Negativ Binomial

$$\sum_1^n Ge(p) = NB(n, p)$$

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

$$\sum_{i=1}^n Po(\lambda_i) = Po(\sum_{i=1}^n \lambda_i)$$

* Summe von Exponential ist Gamma

$$\sum_1^n Exp(\lambda) = Gamma(n, \lambda)$$

* Summe von Normal ist Normal (Summanden müssen nicht mal identisch verteilt sein); vgl. 4.4.1.2.

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

* Weitere Summen (z.B. von X^2) kommen bei den Definitionen von χ^2, F und t-Verteilungen vor.