Stochastik für die Naturwissenschaften

Dr. C.J. Luchsinger

7. $N \to \infty$ (KONVERGENZ: LLN, CLT, POISSONAPPROXIMATION)

"$n \to \infty$" heisst für uns "n gross"
"n gross" heisst aber auch: es kostet Zeit und Geld

Es gehört zur Allgemeinbildung für Akademiker/innen, dass man weiss, was "Gesetz der grossen Zahlen" (LLN) und "Zentraler Grenzwertsatz" (CLT) aussagen!

Anwendung von

LLN: Schätzen von Parametern; siehe auch Kapitel 8, 10 - weil $\overline{X}$ bzw $\overline{x}$ zentral wichtiger Schätzer ist (für den Erwartungswert)

CLT: Approximative Berechnungen, zum Beispiel für Tests; siehe auch Kapitel 8, 9, 10 und weitere

7.1 LLN (Law of Large Numbers; Gesetz der grossen Zahlen)

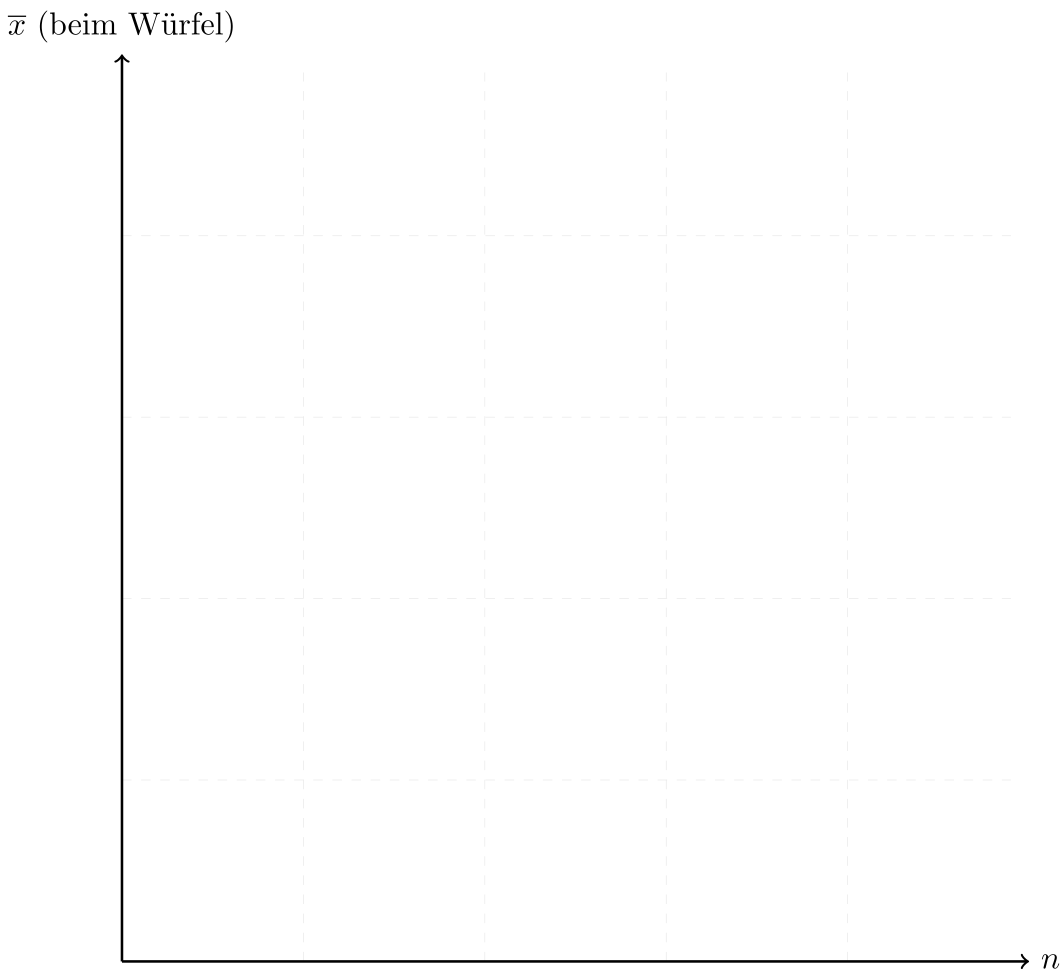
Fragestellung: Was geschieht mit $\overline{X} := \frac{1}{n}\sum_{k=1}^{n} X_k$ falls $n \to \infty$ und welche Voraussetzungen werden wir sinnvollerweise machen? Beispiel fairer Würfel: $\overline{x}$, inwiefern gilt $\overline{x} \to \mu = 3.5$?

Wir haben in Kapitel 2 und in Teil 5.4 über Stichproben immer wieder das arithmetische Mittel von Realisationen (oder Daten) $\overline{x} := \frac{1}{n}\sum_{i=1}^{n} x_i$ (fix, nach Versuch) untersucht. Der Ausdruck $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ (im Büro Luchs, vor Versuch) ist offenbar das Pendant *auf der Ebene der Zufallsgrössen, vor der Realisation* (2 Studis: gleiche Aufgabe mit gleichen, grossen X, danach andere kleine x). $\overline{x}$ konvergiert gegen den Erwartungswert von $\overline{X}$; doch wie genau?

In $\overline{x}$ haben wir reelle Zahlen, da ist kein Zufall mehr drin. Wenn wir aber erneut eine (unabhängige) Stichprobe nehmen, wird $\overline{x}$, zumindest bei stetigen Zufallsgrössen, sicher anders aussehen. Indem wir Erwartungswert und Varianz von $\overline{X}$ berechnen, können wir abschätzen, was mit $\overline{x}$ passieren muss, wenn $n \to \infty$: wir kennen dann die Schwankungsbreite (Varianz, besser Standardabweichung) und wissen, um welchen Wert $\overline{x}$ etwa zu liegen kommt (um den Erwartungswert). Von Kapitel 5, die X_i's seien iid mit $E[X_i] = \mu$ und $V[X_i] = \sigma^2$:

$$E\Big[\frac{1}{n}\sum_{i=1}^{n} X_i\Big] = \mu, \quad V\Big[\frac{1}{n}\sum_{i=1}^{n} X_i\Big] = \frac{1}{n}\sigma^2, \quad \mathrm{sd}\Big[\frac{1}{n}\sum_{i=1}^{n} X_i\Big] = \frac{1}{\sqrt{n}}\sigma.$$

Das heisst, $\overline{x}$ wird um μ herum schwanken, mit immer kleinerer Schwankung, Bild (Tipp für R-Enthusiasten: in R ausprobieren):



Wir wollen eine Aussage der Art machen:

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

konvergiert gegen den Erwartungswert von X_1.

Theorem 7.1 [Satz von Kolmogoroff (LLN)] *Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen, identisch verteilten (sog. iid) Zufallsgrössen mit Erwartungswert μ (und $E[|X_1|] < \infty$ [math]). Dann genügt diese Folge dem Gesetz der grossen Zahlen, das heisst, es gilt für jedes $\epsilon > 0$:*

$$\lim_{n\to\infty} P\left[\ \left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \ < \epsilon\ \right] = 1.$$

Diesen Ausdruck liest man am besten von innen nach aussen!

2 wichtige Anwendungen des LLN

0. Vorbemerkungen

Naturwissenschaft (Labor)/Technik: fairer Würfel: $\bar{x} \to 3.5$ unbestritten da iid.

Andere Gebiete: Noten, Selbstmordraten, Zinsen? iid? Nein, LLN nicht anwendbar.

Klimaerwärmung: auch nicht iid; d.h. LLN nicht anwendbar (gleiche Verteilung ist verletzt (i**i**d) - das ist ja gerade der Punkt: die Temperatur steigt!).

Hingegen darf man $\bar{X}$ und $\bar{x}$ immer berechnen (sobald Intervallskala, zB Durchschnittstemperatur der 1990er Jahre), aber falsche Erwartung, dass $\bar{x} \to \mu$ für irgendein μ.

1. Arithmetisches Mittel

Von der beschreibenden Statistik in Kapitel 2 und Aufgaben in R kennen Sie den Stichprobenmittelwert $\bar{x}$. Dürfen Sie nun den LLN anwenden? Ist die Stichprobe iid (unabhängig, identische Verteilung); wie kann man das wissen?

Wenn das Wort "Stichprobe / Sample" fällt, dann ist das per definitionem immer unabhängig, ausser explizit anders erwähnt (siehe auch 5.4).

iid: Muss a priori die Substanzwissenschaft (GG, Bio) entscheiden - also Sie, nicht ich! A posteriori kann ich als Statistiker dann auch untersuchen, ob es Hinweise gibt, dass die Daten doch nicht unabhängig oder identisch verteilt sind ("Model checking").

Wenn Sie dann iid als Modellannahme akzeptieren, benutzen Sie die Wortwahl: "$\bar{x}$ konvergiert wegen Theorem 7.1 (wegen des LLN, Gesetzes der grossen Zahlen) gegen μ; zB beim Würfel gegen 3.5."

2. Relative Häufigkeiten

Bei den relativen Häufigkeiten gibt es eine kleine Klippe, die wir umschiffen müssen: Theorem 7.1 und obiges Beispiel 1 behandeln ja die Konvergenz gegen einen **Erwartungswert** μ. Wie ist das zu verstehen, wenn man bei relativen Häufigkeiten wie dem Anteil 1er beim Würfel ebenfalls Formulierungen benutzt wie: *wegen des Gesetzes der grossen Zahlen* konvergiert der Anteil 1er gegen 1/6? 1/6 und relative Häufigkeiten sind ja **Wahrscheinlichkeiten** und keine Erwartungswerte?

Um dieses Problem zu lösen müssen wir eine Übersetzungsleistung vollbringen, welche wir mit einem Beispiel begleiten. Wenn Sie 6000 Mal würfeln, dann erwarten Sie (wegen des LLN), dass der Anteil der 1er etwa $1/6 = 1000/6000$ ist.

Dazu führen wir neue Zufallsgrössen Y_i ein und definieren $Y_i = 1$ wenn es eine 1 gibt bei Wurf i und $Y_i = 0$ wenn es keine 1 gibt bei Wurf i. Y_i ist dann natürlich Be(1/6)-verteilt und hat damit Erwartungswert 1/6: $E[Y_i] = p = 1/6$.
Damit ist $\sum_{i=1}^{n} Y_i$ die *Anzahl* 1er; $\frac{1}{n}\sum_{i=1}^{n} Y_i$ ist der *Anteil* 1er. Wegen des LLN gilt nun

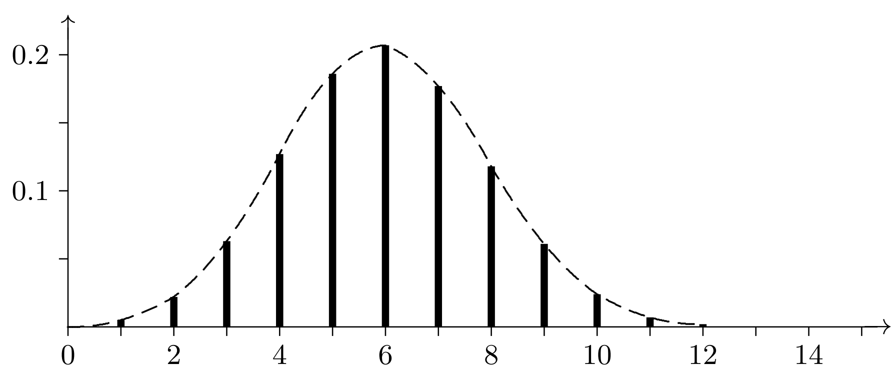
$$\frac{1}{n}\sum_{i=1}^{n} Y_i \to \mu = E[Y_1] = p = 1/6.$$

Nebenbemerkung/Zentral wichtig: Ich kann zum LLN kaum Prüfungsaufgaben stellen, ausser über R-Befehle; vgl. Archiv.

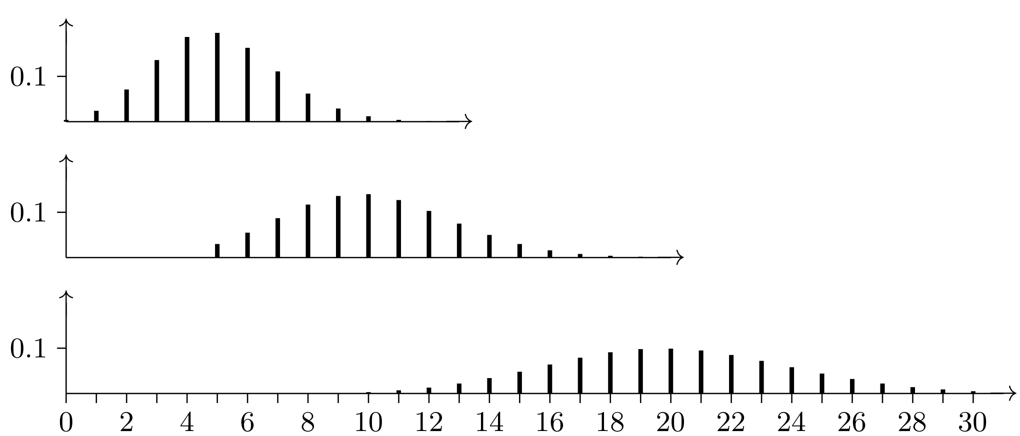
7.2 CLT (Central Limit Theorem; Zentraler Grenzwertsatz)

Sie haben sich sicher schon gefragt, warum immer wieder die Normalverteilung als Modell gewählt wird. Der Grund liegt im Zentraler Grenzwertsatz CLT. Schauen wir das deshalb genauer an.

Zeichnet man das Stabdiagramm einer Binomialverteilung für grössere Werte von n, und verbindet man die Enden der “Stäbe”, so gewinnt man den Eindruck, dass ungefähr eine Glockenkurve entstanden sei. Dieser Eindruck wird umso besser, je grösser n ist.



Allerdings verschiebt sich diese Kurve mit wachsendem n (und bei festgehaltenem p) immer mehr nach rechts, was nicht überrascht, da der Erwartungswert einer binomial verteilten Zufallsgrösse $= np$ ist (vgl. Kapitel 5.2).



Wir stellen zusammengefasst vier Dinge fest:

1. Wir haben eine *Summe* von iid Zufallsgrössen angeschaut (Summe von iid Bernoulli ist binomial). Auch bei der Summe von iid U[0,1], siehe https://de.m.wikipedia.org/wiki/Datei:Summe-von-Gleichverteilungen3.svg, zeigt sich das gleiche Bild!
2. Die Verteilung wanderte immer (in diesen Beispielen) nach rechts.
3. Die Varianz der Verteilung wurde immer grösser.
4. Am Schluss sah es aus wie eine Normalverteilung (das *ist* ein **Weltwunder**: in der Mathematik beweisen wir das zwar - verstehen tun wir es auch nicht).

Der erste Punkt (das mit der Summe) ist notwendig; die Idee dahinter ist, dass sich Effekte der (unabhängigen!) Summanden ausgleichen. Wir betrachten also erstmal eine Summe $\sum_{k=1}^n X_k$. Der zweite und dritte Punkt sind eher störend, weil wir das Ganze am Schluss mit einer *Standard*normalverteilung vergleichen wollen. Deshalb zentrieren und normieren wir die Summe und untersuchen stattdessen mit $\mu := E[X_k], \sigma^2 := V[X_k]$:

$$\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma}$$

Woran erinnert Sie das?

Überprüfen Sie, dass jetzt der Erwartungswert 0 ist und die Varianz gleich 1 (vgl Rechnungen bei Z-Transformation in Kapitel 5).

$E\left[\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma}\right] =$

$V\left[\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma}\right] =$

Deshalb müssen Sie die Gesetze für $E[.]$ und $V[.]$ kennen!

Theorem 7.2 [Zentraler Grenzwertsatz] *Sei $X_k, k \geq 1$, eine Folge von iid Zufallsgrössen mit $E[X_1] =: \mu$ und $0 < V[X_1] =: \sigma^2 < \infty$. Dann gilt für $a \in \mathbb{R}$ beliebig:*

$$P\left[\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \leq a\right] \longrightarrow P[\mathcal{N}(0,1) \leq a]$$

wenn $n \to \infty$.

oft Prüfungsaufgabe 4, oft mit R versetzt, siehe Archiv.

Bsp 1, zusammen: Wie gross ist die Wahrscheinlichkeit, dass bei 100 Würfen mit einer fairen Münze mindestens 65 mal Kopf erscheint? **KLICKER**

Bsp 1, nochmals mit **Diskretisierungskorrektur** (englisch Continuity Correction); Diskretisierungskorrektur im FS 2025 kein Prüfungsstoff; ab 2026 Prüfungsstoff.

Wie gross ist die Wahrscheinlichkeit, dass bei 100 Würfen mit einer fairen Münze mindestens 65 mal Kopf kommt?

Exakt wäre mit R:

Bild zu Diskretisierungskorrektur

Bsp 2, alleine: Sei $\lambda = 2$ und $(X_k)_{k=1}^{50}$ iid $\exp(\lambda)$. Berechnen Sie $P[\sum_{k=1}^{50} X_k > 22]$.

Hausaufgabe: nächste 4 Seiten lesen

Bemerkungen: 1. Ab welchem n kann/darf/muss man die Normalverteilung brauchen? Nichttriviale, allgemeine Fehler-Abschätzungen, welche nur von n abhängen und für alle Verteilungen (der X_k) gelten, existieren nicht. Sicher kommt es auch auf die Anwendung an. Wenn die X_k's aber Be(p)-Zufallsgrössen sind, so sagt eine Bauernregel, dass $np(1-p) \geq 9$ gelten sollte, damit man die Normalverteilung benutzen darf. Wenn nicht anders erwähnt, können Sie in meiner Vorlesung, Übungen und Prüfung diese Regel ab jetzt anwenden.

2. Was ist der Grund für den häufigen Einsatz der Normalverteilung? Die gemessenen Grössen sind oft Summen von kleinen Effekten ("$\sum_k X_k$"). Dann kommt der CLT zum Zug. Dabei ist Unabhängigkeit wichtiger als die gleiche Verteilung.

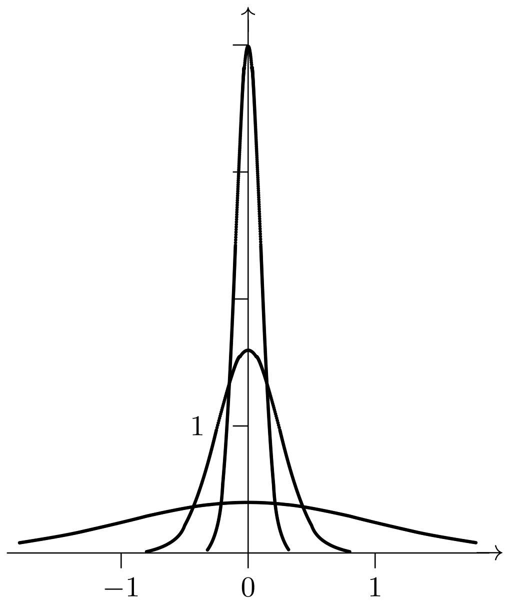
3. Wenn $X \sim \mathcal{N}(\mu, \sigma^2)$, dann gilt (siehe Kapitel 5) $P[X \in [\mu \pm 2\sigma]] \doteq 0.95$. Wegen des CLT gelten deswegen folgende Praktikerregeln auch für andere Verteilungen:

* $Y \sim \text{Bin}(n,p) \Rightarrow P[Y \in [np \pm 2\sqrt{np(1-p)}]] \doteq 0.95$ wenn $np(1-p) \geq 9$.

* $Z \sim \text{Po}(\lambda) \Rightarrow P[Z \in [\lambda \pm 2\sqrt{\lambda}]] \doteq 0.95$ wenn λ gross.

4. Beim LLN teilen wir die Summe durch n und $\bar{X}$ konvergiert dann gegen einen Punkt μ. Beim CLT teilen wir nur noch durch $\sigma\sqrt{n}$ und damit bleibt die Varianz konstant gleich 1, egal wie n (siehe Berechnung vor 3 Seiten).

5.1. $\bar{X}$: Falls die X_i schon normalverteilt sind, dann ist auch die iid Summe normalverteilt (Schluss von Kapitel 6) und wenn wir noch durch n teilen bleibt es normalverteilt (vgl Z-Transform). Insgesamt haben wir $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$; vgl Kapitel 5. Das hat bis jetzt noch nichts mit dem CLT zu tun! Bilder zu den Dichten von $\overline{X}$ mit $n = 1, 16, 100$:



5.2. Wegen des CLT brauchen wir dann aber bei grossem n nicht mehr zwingend, dass die X_i schon normalverteilt sind. Auch wenn wir mit anderen Verteilungen starten (Uniform, bimodal) ist $\bar{X} \dot{\sim} \mathcal{N}(\mu, \sigma^2/n)$. Auf Übungsblatt 10 (nicht jedes Jahr) betrachten wir dazu das Video www.youtube.com/watch?v=jvoxEYmQHNM . Vorsicht wegen der dortigen Terminologie: Dort werden zwei Schritte zusammengemischt: der CLT macht lediglich eine Aussage über die (standardisierte) Summe. Erst in diesem 5. Punkt folgt dann die Untersuchung von $\bar{X}$.

Vorsicht, Verwechslungsgefahren:

Erstmal zu einer einzelnen Zufallsgrösse X:

1. Es gibt die ganz normale Z-Transformation: Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann gilt:

$$\frac{X - \mu}{\sigma}$$

ist $\mathcal{N}(0, 1)$.

2. Ich kann und will Ihnen aber nicht verbieten, dass Sie bei einem Y mit $E[Y] = \mu, V[Y] = \sigma^2$, aber Y sei explizit *nicht* normalverteilt, auch folgende Transformation machen:

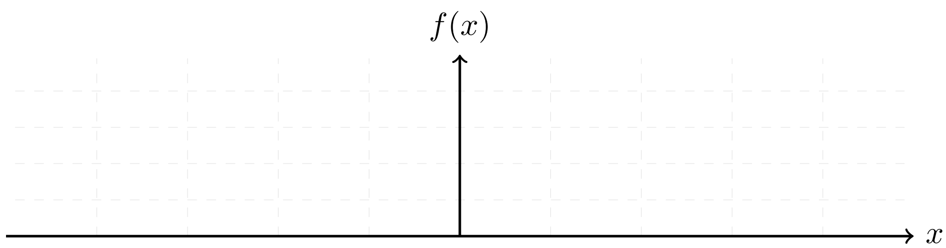
$$Z := \frac{Y - \mu}{\sigma}.$$

Für ein solches Z gilt: $E[Z] = 0, V[Z] = 1$ - aber Z ist nach wie vor *nicht* normalverteilt. Man hat einfach zentriert und normiert (zusammen: standardisiert).

3. Dieses *standardisieren* kann man auch mit Daten machen (wieder ist Normalverteilung nicht zwingend): der R-Befehl lautet einfach

f <- (e-mean(e))/sd(e).

Egal, mit was Sie in e starten, f ist danach standardisiert. Der Vorteil einer Standardisierung ist der, dass die *Form* verschiedener Verteilungen (zB die Dichten) besser miteinander verglichen werden kann:



Dann zu einer Summe von Zufallsgrössen $\sum_{k=1}^{n} X_k$:

4. Zum Schluss noch die Betrachtung einer Summe von (nicht zwingend normalverteilter) Zufallsgrössen: auch hier machen wir so etwas wie eine Z-Transform der gesamten Summe: wir betrachten

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}.$$

Dieser Ausdruck konvergiert mit $n \to \infty$ gegen eine $\mathcal{N}(0,1)$-Zufallsgrösse. Das ist der CLT!

5. Zu guter Letzt - immer wieder in Übungen und Prüfungen anzutreffen - sei noch erwähnt, dass wenn im CLT die X_k's bereits normalverteilt sind, dann ist

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$$

für jedes n schon exakt (!) normalverteilt (also auch schon für endliches n).

Billige Kritik an der Normalverteilung und die Entgegnungen:

Merken Sie sich im Umgang mit "Wissen" von "Experten": wenn es um die Frage geht, ob etwas für die Praxis relevant ist oder nur blosse Theorie, dann entscheiden viele nach der Regel: *"Alles was ich noch verstehe, ist wichtig für die Praxis - und alles was ich nicht mehr verstehe, ist blosse Theorie und unwichtig in der Praxis".*

Eine Kritik in dem Sinne könnte sein: "Normalverteilung bei Gewicht oder Länge? Aber Gewicht und Länge können doch in der Praxis unmöglich negativ sein!"

rhetorisch gute Volte: erstmal Recht geben: "Ja, stimmt: die Normalverteilung geht von $-\infty$ bis ∞."

dann Gegenangriff:

1. Dichtefunktion $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ geht *sehr schnell* gegen 0. Das gilt für Exponentialfunktionen generell - hier kommt das x noch im Quadrat vor.

2. Passt wegen CLT sehr gut zu sehr vielem; grosse Ausnahme Riskmanagement in Banken.

3. Mathematik funktioniert mit Normalverteilung sehr gut.

7.3 Zusammenhänge bei Verteilungen II: Limesverteilungen, vgl 6.3

In "6.3 Zusammenhänge bei Verteilungen I", haben wir uns gefragt, wie Summen von unabhängigen Zufallsgrössen verteilt sind. Jetzt geben wir eine Verteilung an (Binomial), welche zu einer anderen Verteilung wird (Poisson), wenn wir *gezielt* einen Parameter gegen ∞ gehen lassen.

Dazu betrachten wir erstmal, vermeintlich ohne Zusammenhang, eine *Folge* von Binomialverteilungen. Wir schauen in unterer Tabelle jeweils die Wahrscheinlichkeit $P[X = 0]$ bzw. $P[X = 2]$ an. Die Folge von Binomialverteilungen hat aber eines gemeinsam: der Erwartungswert ist immer 2. Der Erwartungswert ist bekanntlich np; wenn er konstant bleibt, muss p kleiner werden, wenn n grösser wird. Man kann sich fragen, ob in einer Zeile der unteren Tabelle der Wert, zum Beispiel für $P[X = 0]$, konvergiert, wenn $n \to \infty$. Dies ist überraschenderweise so. Der Wert konvergiert gegen $P[Y = 0]$, wenn Y eine Poisson-Zufallsgrösse ist mit Parameter $\lambda = 2$.

	Bin(50,0.04)	Bin(100,0.02)	Bin(200,0.01)	Bin(1000,0.002)
$P[X = 0]$	0.1299	0.1326	0.1340	0.1351
$P[X = 2]$	0.2762	0.2734	0.2720	0.2709
np	2	2	2	2

Wenn man die Wahrscheinlichkeitsfunktion (die Werte von $P[X = k], k \geq 0$) einer Binomial- und einer Poissonverteilung mit gleichem Erwartungswert vergleicht, so sieht man, dass der Unterschied immer kleiner wird, je gösser n in der Binomial-Verteilung ist. Zum Vergleich die Werte der Wahrscheinlichkeitsfunktion von Bin(200, 0.01) und Po(2) bei $k = 0, 1, 2, 3$.

k	0	1	2	3
$\mathrm{Bin}(200, 0.01)$	0.1339797	0.270666	0.272033	0.181355
$\mathrm{Po}(\lambda)$, $\lambda = 2$	0.1353353	0.270671	0.270671	0.180447

Im folgenden Satz ist λ konstant. p ist klein und wird immer kleiner; deshalb indexieren wir es mit einem n: p_n. Wir haben in unterem Satz eine ganze Folge von Binomialverteilungen. Dies ist für die Formulierung des Satzes wichtig, nicht aber für die Anwendungen. Für die Anwendung ist folgender Punkt wichtig: Weil der Erwartungswert einer $\mathrm{Po}(\lambda)$-Zufallsgrösse gleich λ ist und der Erwartungswert einer $\mathrm{Bin}(n, p)$ gleich np, fordern wir (damit sicher mal die Erwartungswerte übereinstimmen) $np = \lambda$, oder eben $np_n = \lambda$.

Satz 7.3 [Konvergenz der Binomial- gegen die Poisson-Verteilung] *Sei X_n eine Folge von $Bin(n, p_n)$-Zufallsgrössen. Es gelte für alle n: $np_n = \lambda > 0$ (Erwartungswerte jeweils gleich). λ ist konstant! Dann gilt für $k \in \mathbb{N}_0$ fest:*

$$P[X_n = k] \longrightarrow e^{-\lambda}\frac{\lambda^k}{k!}$$

für $n \to \infty$. Damit haben wir im Limes eine Poisson-Verteilung mit Parameter λ.

Vorteil: einfachere Formel bei Poisson als bei Binomial (Binomialkoeffizient!).

Wie wird das ganze in der Praxis eingesetzt (identischer Text in Kapitel 6, siehe auch 6.1.4)?

1. Situation, wo erste Idee $\mathrm{Bin}(n, p)$, zB # Hausbrände in CH pro Jahr, # Meteoriteneinschläge pro Jahr, # Flugzeugabstürze pro Jahr
2. n gross und p klein
3. Satz 7.3: nehme stattdessen $\mathrm{Po}(\lambda)$, da einfacher und nur 1 Parameter statt 2
4. λ so, dass $E[\mathrm{Bin}(n, p)] = E[\mathrm{Po}(\lambda)]$, das heisst $np = \lambda$ (manchmal kennt man eher λ und nicht np (Meteoriteneinschläge))

1. Beispiel:

2. Beispiel:

Schiffsunglücke oder Flugzeugabstürze von Airlines ("If God had wanted man to fly, he would have given him wings!")

Die Anzahl Abstürze von Airlines pro Jahr (n gross, p klein) wird mit einer Poisson-Zufallsgrösse modelliert, mit unterschiedlichem λ von Airline zu Airline. Bei doppelt so vielen Flügen gibt es tendenziell ein doppeltes λ. Tipp bei der Auswahl der Airline: auf die generelle Sicherheitskultur eines Landes achten und die Airline entsprechend wählen. Dabei ist die Swiss(air) leider eine grosse Negativausnahme. Die Gründe waren aber meist Terrorismus oder technische Fehler, welche nicht der Swiss(air) angelastet werden können. Zudem war die Swiss(air) im Verhältnis zur Bevölkerung eine sehr grosse Airline. Mehr auf (nicht clickbarer Link, abschreiben):

`https://de.wikipedia.org/wiki/Liste_von_Flugunfällen_(Schweiz)`

PrüfungsTAKTIK - wann Bin, Po, $\mathcal{N}$? [strenge MATHE vs TAKTIK in der Prüfung - Prüfungsaufgaben sind aber immer klar formuliert, siehe Archiv]

* Wenn es heisst: Sei $X \sim Bin(n,p)$ oder $X \sim Po(\lambda)$ oder $X \sim \mathcal{N}(\mu, \sigma^2)$: einfach lösen

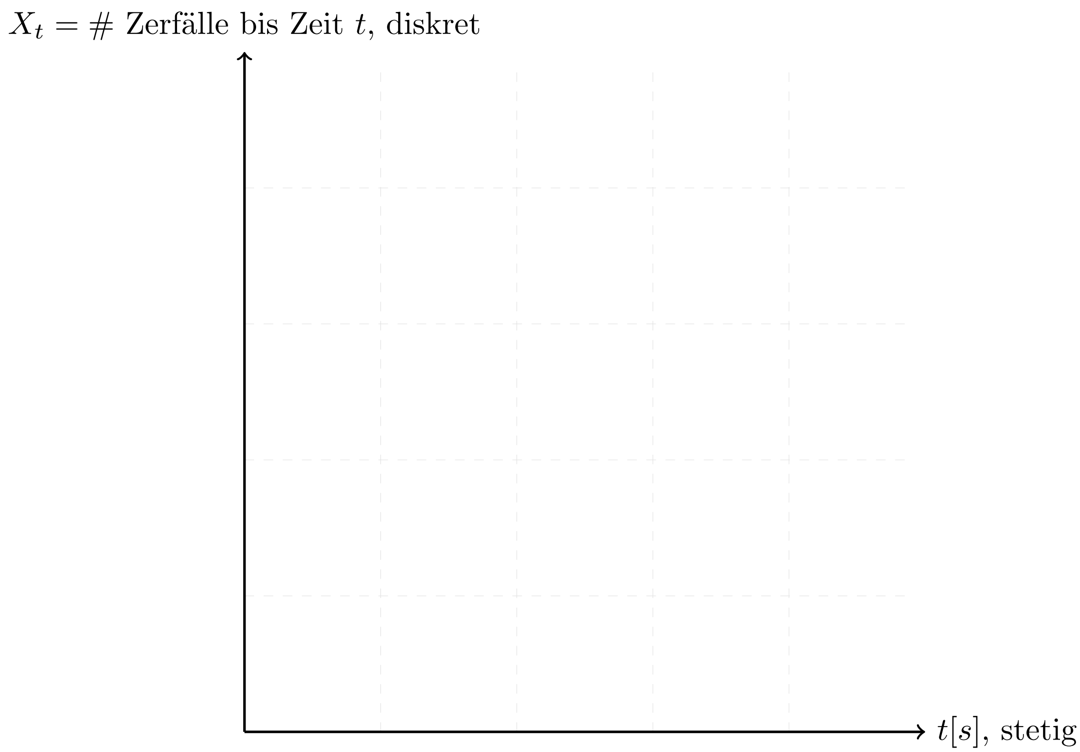
* Bin mit Po ersetzen, wenn: n gross und p klein (Satz 7.3) [MATHE] und auch wenn Frage nach $P[X = k]$ für wenige Werte von k (1-4 Werte, n gross) [TAKTIK]

* Bin mit $\mathcal{N}$ ersetzen, wenn: $np(1-p) \geq 9$ (CLT) [MATHE] und auch wenn Frage nach $P[X \leq a]$ oder $P[X \in [168, 188]]$ - langes Intervall [TAKTIK]

Exkurs Poisson**prozess**: Im Jahr 1910 zählten Rutherford und Geiger während 326 Minuten die Zerfälle bei einem radioaktiven Poloniumpräparat. Die Zeitspanne wurde in Intervalle von 7.5 Sekunden Länge unterteilt. In der folgenden Tabelle ist die Anzahl Intervalle angegeben, in denen 0, 1, 2, 3.... Zerfälle erfolgten:

# Zerfälle	0	1	2	3	4	5	6	7	8	9	10	11	≥ 12
# Intervalle	57	203	383	525	532	408	273	139	45	27	10	4	2

Wie ist das zu verstehen? Beachten Sie auch, wie (unerwartet) die Exponentialverteilung und die Poissonverteilung mit einander zu tun haben.



Wir hatten bei der Einführung der Exponentialverteilung festgehalten, dass diese Verteilung (wegen der Gedächtnislosigkeit sowohl der Exponentialverteilung wie auch des Isotops) *das* Modell für den Zerfall eines *einzelnen* Atoms ist. In obiger Skizze wurde jetzt aber ein (im Verlauf der Zeit immer kleiner werdender) *ganzer Klumpen* eines Isotops untersucht und wieder festgehalten, dass wir eine exponentialverteilte Zwischenzeit haben, bis wieder ein Isotop zerfällt. Dies ist wirklich überraschend, oder etwa doch nicht?

7.4 Mind-Mapping Verteilungen

Wichtig:

1. Lesen Sie jetzt das komplette Kapitel im Buch selber durch.
2. Lösen Sie danach mindestens 5 Aufgaben hinten im Kapitel und vergleichen Sie mit den Lösungen am Schluss des Buches. Bei Bedarf lösen Sie mehr Aufgaben.
3. Gehen Sie in die Übungsstunde. Drucken Sie das Übungsblatt dazu *vorher* aus, lesen Sie *vorher* die Aufgaben durch und machen sich erste Gedanken dazu (zum Beispiel, wie man sie lösen könnte).
4. Dann lösen Sie das Übungsblatt: zuerst immer selber probieren, falls nicht geht: Tipp von Mitstudi benutzen, falls immer noch nicht geht: Lösung von Mitstudi anschauen, 1 Stunde warten, versuchen, aus dem Kopf heraus wieder zu lösen, falls immer noch nicht geht: Lösung von Mitstudi abschreiben (und verstehen - also sollte man insbesondere keine Fehler abschreiben!).
5. Lösen Sie die entsprechenden Prüfungsaufgaben im Archiv.

Lessons learnt:

In 3Blue1Brown
https://www.3blue1brown.com/lessons/clt
https://www.3blue1brown.com/lessons/gaussian-convolution

LLN
$$\lim_{n\to\infty} P\left[\ \left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \epsilon\ \right] = 1.$$

CLT
$$P\left[\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le a\right] \longrightarrow P[\mathcal{N}(0,1) \le a]$$

CLT bei Binomial ab $np(1-p) > 9$.

$Y \sim \mathrm{Bin}(n,p) \Rightarrow P[Y \in [np \pm 2\sqrt{np(1-p)}]] \doteq 0.95$ wenn $np(1-p) \ge 9$.

$Z \sim \mathrm{Po}(\lambda) \Rightarrow P[Z \in [\lambda \pm 2\sqrt{\lambda}]] \doteq 0.95$ wenn λ gross.

Binomial und n gross, p klein; $\lambda := np$:
$$P[X_n = k] \longrightarrow e^{-\lambda}\frac{\lambda^k}{k!}$$

Klickerfragen zum Aufwärmen:

Frage 1: Was beschreibt das Gesetz der grossen Zahlen (LLN)? (a) Die Summe unabhängiger Zufallsgrössen folgt immer einer Normalverteilung. (b) Der Mittelwert einer Stichprobe konvergiert bei wachsendem Stichprobenumfang gegen den Ewartungswert. (c) Die Wahrscheinlichkeit für ein Ereignis wird mit der Zeit immer kleiner.

Frage 2: Welche Bedingung muss für das LLN erfüllt sein, wenn der Erwartungswert existiert (a) Die Zufallsgrössen müssen normalverteilt sein. (b) Die Zufallsgrössen müssen stetig sein (c) Die Zufallsgrössen müssen unabhängig und identisch verteilt (iid) sein.

Frage 3: Man würfelt 10.000-mal mit einem fairen Würfel. Was passiert langfristig mit dem Mittelwert der Augenzahlen? (a) Er nähert sich 3.5. (b) Er schwankt immer stärker. (c) Er bleibt bei etwa 6.

Frage 4: Was besagt der Zentrale Grenzwertsatz (CLT)? (a) Der grösste Wert einer Stichprobe wird immer kleiner. (b) Die Wahrscheinlichkeit für seltene Ereignisse wird immer höher. (c) Die Summe vieler unabhängiger Zufallsgrössen nähert sich nach Standardisierung einer Normalverteilung an.

Frage 5: Welche Bedingung ist wichtig für den Zentralen Grenzwertsatz? (a) Die Zufallsgrössen müssen normalverteilt sein. (b) Die Zufallsgrössen müssen unabhängig und identisch verteilt (iid) sein. Zudem muss die Varianz existieren. (c) Die Stichprobengrössen müssen kleiner als ein beliebiges $a \in \mathbb{R}$.

Frage 6: Wann kann man eine Binomialverteilung durch eine Poissonverteilung approximieren? (a) Wenn n sehr klein ist. (b) Wenn p genau 0.5 beträgt. (c) Wenn die Erfolgswahrscheinlichkeit p sehr klein und n sehr gross ist.

Frage 7: Welcher Zusammenhang muss zwischen n und p bestehen, damit die Poisson-Approximation funktioniert? (a) $np = \lambda$. (b) $n = p$. (c) $n = \sqrt{p}$

Frage 8: Warum ist das LLN bei Klimadaten problematisch? (a) Weil die Temperaturen nicht unabhängig und identisch verteilt sind. (b) Weil es keine Erwartungswerte gibt. (c) Weil Wetterdaten immer normalverteilt sind.

Frage 9: Du wirfst eine faire Münze 1000-mal. Welcher Anteil an Kopf wird erwartet? (a) Exakt 50 %, (b) Etwa 50%, (c) Etwa 30%.

Frage 10: Was passiert mit der Schwankungsbreite des Mittelwerts, wenn der Stichprobenumfang n grösser wird? (a) Sie wird grösser. (b) Sie bleibt gleich. (c) Sie wird kleiner.

Frage 11: In einer Stadt gibt es pro Jahr durchschnittlich 3 Erdbeben. Welche Verteilung eignet sich zur modellierung der Anzahl Erdbeben? (a) Normalverteilung. (b) Poisson-Verteilung. (c) Exponentialverteilung

Frage 12: Ab welcher Bedingung kann man eine Binomialverteilung durch die Normalverteilung annähern? (a) Wenn $np(1-p) \geq 9$, (b) Wenn $n = 100$, (c) Wenn $p = 0.5$

Lösungen zu Klickerfragen:

Frage 1: (b) Der Mittelwert einer Stichprobe konvergiert bei wachsendem Stichprobenumfang gegen den Erwartungswert; Frage 2: (c) Die Zufallsgrössen müssen unabhängig und identisch verteilt (iid) sein; Frage 3: (a) Er nähert sich 3.5; Frage 4: (c) Die Summe vieler unabhängiger Zufallsgrössen nähert sich nach Standardisierung einer Normalverteilung an; Frage 5: (b) Die Zufallsgrössen müssen unabhängig und identisch verteilt (iid) sein. Zudem muss die Varianz existieren; Frage 6: (a) Wenn die Erfolgswahrscheinlichkeit p sehr klein und n sehr gross ist; Frage 7: (a) $np = \lambda$; Frage 8: (a) Weil die Temperaturen nicht unabhängig und identisch verteilt sind; Frage 9: (b) Etwa 50%; Frage 10: (c) Sie wird kleiner; Frage 11: (b) Poisson-Verteilung; Frage 12: (a) Wenn $np(1-p) \geq 9$.