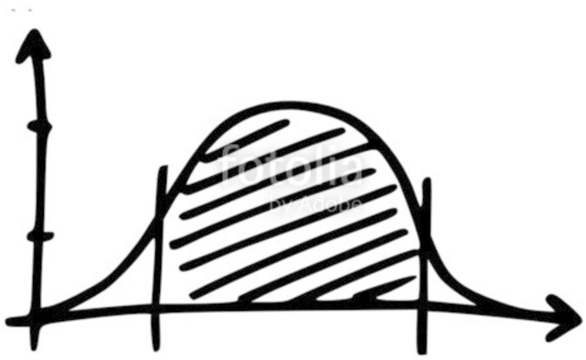
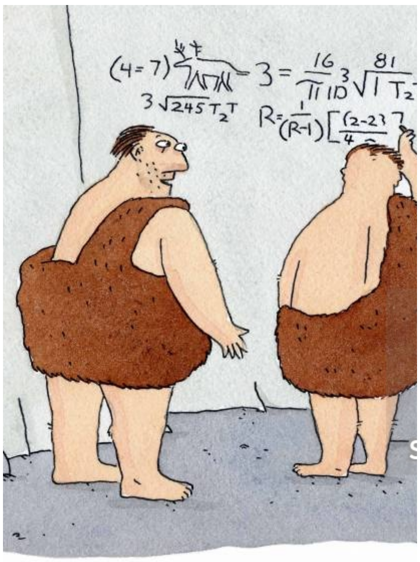
Datum Name

Wahrscheinlichkeits-verteilungen

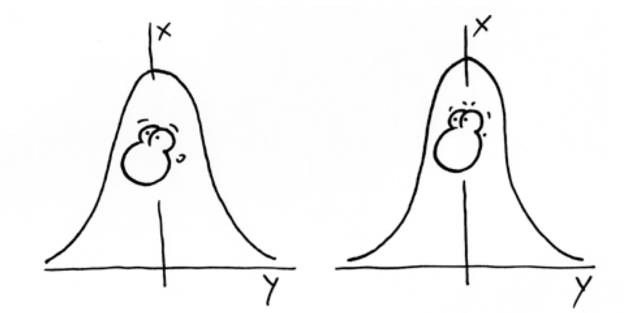


beschreibende Statistik
Verteilungen
Binomial-, Normal-, Poisson-Verteilung



“I like the part about the deer.”

Inhalt



"I always feel so normal, so bored, you know. Sometimes I would like to do something... you know... something... mmm... Poissonian."

0 Rückblick – Beschreibende Statistik

Merkmal, Umfang, Verteilung, absolute Häufigkeit, relative Häufigkeit, Histogramm

Beispiel 1 Noten

Noten in einer Klasse von 10 Schülern. (Der „**Umfang**“ beträgt also n = 10.)

Verteilung

Note (**Merkmal**)	3	4	5	6
absolute Häufigkeit H	II	III	IIII	I
relative Häufigkeit h	0.2	0.3	0.4	0.1

a) Warum ist die relative Häufigkeit h „interessanter“ (wichtiger) als die absolute Häufigkeit H?

b) Zeichnen Sie ein **Histogramm** dieser Tabelle (mit Säulenbreite $\Delta = 1$).
Wie gross ist die „Fläche“? (Wo sind Sie „Flächen“ und „Säulen“ schon begegnet?)

c) Die Note ist ein **Merkmal** eines Schülers.
Ein Merkmal X kann verschiedene Werte x_1, x_2, x_3, ... annehmen.
Im Beispiel nimmt das Merkmal

X = Note eines Schülers der Klasse

die Werte an:

$x_1 = 3$, $x_2 = 4$, $x_3 = 5$, und $x_4 = 6$.

Welche Werte nimmt wohl das folgende Merkmal X an?

X = Schuhgrösse eines Schülers der Klasse

Mittelwert (Durchschnitt)

Beispiel 2 Notenschnitt

a) Welchen Notenschnitt hat die Klasse in Beispiel 1?

b) Der Mittelwert lässt sich mit Hilfe der absoluten bzw. relativen Häufigkeiten wie folgt berechnen:

$$\textbf{Mittelwert } \bar{x} = \frac{1}{n} \sum_{i=1}^{k} x_i \cdot H(x_i) = \sum_{i=1}^{k} x_i \cdot h(x_i)$$

Bestätigen Sie diese Formeln!

c) In einer Klasse haben 20% die Note 4.5, 35% die Note 5, 30% die Note 5.5 und der Rest die Note 6.

- Berechnen Sie den Notendurchschnitt.
- Muss man – um den Durchschnitt zu berechnen – nicht die Gesamtanzahl Schüler der Klasse wissen?

d*) Schreiben Sie die Formel für den Mittelwert hin.

- Man nennt die $h(x_i)$ auch *Gewichte* der Werte x_i. Warum?
- Was bedeutet n und was bedeutet k in den Formeln? Gibt es einen Zusammenhang zwischen n und k?

Streuung (Abweichung)

Beispiel 3 **Wie misst man Streuung?**

Gegeben sind vier Datensätze vom Umfang 3

I 2;2;2 **II** 1;2;3 **III** 0;2;4 **IV** 1;1;4

a) Welche Daten streuen am meisten? Warum?

b) Eine naheliegende Idee, die Streuung zu messen, besteht darin, die
durchschnittliche Abweichung d vom Mittelwert
zu berechnen. Tun Sie dies für die vier Datensätze.

Beispiel 4 **Varianz, Standardabweichung**

a) Das gebräuchlichste („Standard") Streuungsmass ist die sogenannte Standardabweichung.
Sie ist definiert, wie folgt:

$$\textbf{Standardabweichung } s = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot H(x_i)}$$

Versuchen Sie, den „Aufbau" dieser Formel zu erklären.

b) Wie lautet die Formel der Standardabweichung, geschrieben mit relativen Häufigkeiten?

c*) Zeigen Sie an einem Beispiel, dass gilt: $d \neq s$.

Zusammenfassende Aufgabe

Beispiel 5 **Geschwister & Autos**

a) In Ihrer Klasse wird das Merkmal X = Anzahl Geschwister untersucht.

- Erstellen Sie eine Häufigkeitstabelle und zeichnen Sie ein Histogramm.
- Berechnen* und interpretieren Sie den Mittelwert $\bar{x}$ und die Standardabweichung s.

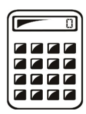
b) In der Schweiz haben 40% aller Haushalte ein Auto, 30% haben zwei, 15% drei und der Rest keines.

- Erstellen Sie eine Häufigkeitstabelle und zeichnen Sie ein Histogramm für X = Anzahl Autos.
- Berechnen* und interpretieren Sie den Mittelwert $\bar{x}$ und die Standardabweichung s.

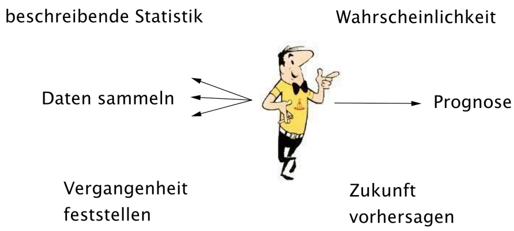
Taschenrechner

Folgende Tasten sind wichtig:

- **data** (Eingeben von Werten x_i in Liste 1; zugehörige Häufigkeiten in Liste 2)
- **stat** (Auswerten)
- **data 2-mal** (Löschen)



1 Einstieg



Beispiel 6 **relative Häufigkeit h vs Wahrscheinlichkeit p**

Versuchen Sie Unterschied und Gemeinsamkeit zu benennen von

- relativer Häufigkeit h
- Wahrscheinlichkeit p.

Sämtliche Konzepte werden von der beschreibenden Statistik auf die „Wahrscheinlichkeitstheorie" übertragen.

Die Tabelle zeigt die Verwandtschaft einander entsprechender Begriffe.

Statistik	**Wahrscheinlichkeitstheorie**
Bei einem **Merkmal X** wird die **relative Häufigkeit h_i** eines Wertes x_i **„praktisch"** bestimmt mit „**Zählen**"	Bei einer **Zufallsgrösse X** wird die **Wahrscheinlichkeit p_i** eines Wertes x_i **„theoretisch"** bestimmt mit „**Überlegen** ..."
Damit ergeben sich direkt	
Häufigkeitsverteilung (Tabelle, Histogramm)	**Wahrscheinlichkeitsverteilung (Tabelle, Histogramm)**
Mittelwert $\bar{x}$	**Erwartungswert μ**
Standardabweichung s	**Standardabweichung σ**

Beispiel 7 **1 Würfel**

Ein Würfel wird geworfen. Wir betrachten (*definieren!)* die Zufallsgrösse

X = Anzahl Augen

a) Geben Sie die **Wahrscheinlichkeitsverteilung von X** an (immer: Tabelle und Histogramm).

b) Schätzen, berechnen und interpretieren Sie den Erwartungswert μ und die Standardabweichung σ.

c) Geben Sie die **allgemeinen Formeln Erwartungswert μ und die Standardabweichung σ** an!

Bei den folgenden Aufgaben empfiehlt sich folgendes Vorgehen:

- zuerst überlegen, welche Werte x_i die Zufallsgrösse X überhaupt annehmen kann
- diese Werte in der Verteilungstabelle notieren
- evt. in einer „Nebenrechnung" die zugehörigen Wahrscheinlichkeiten p_i berechnen

Beispiel 8 **2 Würfel**

a) Wir *definieren* die Zufallsgrösse

X = Summe der Augen

- Geben Sie die Wahrscheinlichkeitsverteilung von X an
- Schätzen, berechnen und interpretieren Sie μ und σ.

b) Wie **a)**. Wir *definieren* die Zufallsgrösse

Y = grösste Zahl

Beachte *beim gleichen Zufallsexperiment lassen sich verschiedene Zufallsgrössen betrachten (definieren).*

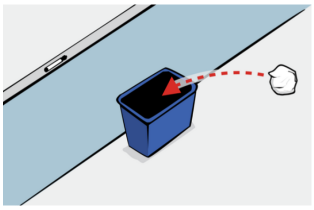
Beispiel 9 **Spiel**

Sie dürfen drei Würfel werfen. Erscheint die „6“ nicht, dann müssen Sie 50 Rp. bezahlen.
Erscheint die „6“ einmal, zweimal oder sogar dreimal, dann gewinnen Sie 1 Fr., bzw. 2 Fr., bzw. 3 Fr.
Spielen Sie?
Wir betrachten dazu die Zufallsgrösse

X = Gewinn (bei einem Wurf mit drei Würfeln)

Berechnen und interpretieren Sie μ und σ.

Beispiel 10 **vermischte**



a) Bucher trifft den Güselkübel mit einer Wahrscheinlichkeit von p = 40%.
Er wirft immer höchstens 4-mal; dann gibt er auf...
Wir *definieren* die Zufallsgrösse

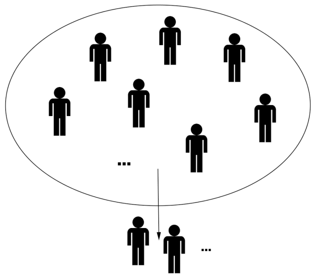
T = Anzahl Versuche

- Geben Sie die Wahrscheinlichkeitsverteilung von T an.
- Berechnen und interpretieren Sie Erwartungswert und Standardabweichung von T.

b) Eine Urne enthält 4 rote und 3 weisse Kugeln. 2 Kugeln werden nacheinander ohne Zurücklegen gezogen.
X sei die Anzahl roter Kugeln.
Bestimmen Sie die Verteilung von X und berechnen Sie $\mu(X)$ und $\sigma(X)$.

c) Eine Zufallsvariable X nimmt den Wert 1 mit der Wahrscheinlichkeit p an und den Wert 0 mit der Wahrscheinlichkeit 1 – p. Berechnen Sie den Erwartungswert $\mu(X)$ und die Standardabweichung $\sigma(X)$.

Beispiel 11 **Stichprobe**



Der SVP Wähleranteil beträgt rund 30%. Man wählt nun „zufällig“ eine Gruppe von Personen aus, d.h. man zieht eine **Stichprobe**.
Sei X = Anzahl SVP Wähler in der erhobenen Stichprobe.
Geben Sie die Verteilung von X an und berechnen und interpretieren Sie μ und σ, wenn die Stichprobe

a) 3 Personen umfasst

b) 10 Personen umfasst

Beachte *Offensichtlich ist b) in der Berechnung eher mühsam. Das ruft nach einem Muster...*

2 Verteilungen

Beim gleichen Zufallsexperiment kann man das Augenmerk auf ganz verschiedene Aspekte legen.
Mathematisch formuliert:
Beim gleichen Zufallsexperiment kann man die Zufallsgrösse X ganz verschieden *definieren*.

Beispiel 12 verschiedene Zufallsgrössen bei demselben „Zufallsversuch“ (Experiment)

Definieren Sie mindestens zwei verschiedene Zufallsgrössen für das Zufallsexperiment:

- ein Würfel wird 3-mal geworfen.
- eine Person wird aus der Klasse ausgewählt.

Es gibt jedoch Zufallsgrössen, die immer wieder auftreten und ihre Verteilungen weisen ein spezielles Muster auf. Dieses Muster fassen wir in eine „Formel“ und können so das mühsame Aufstellen einer Tabelle umgehen.

Weil diese „speziellen“ Verteilungen immer wieder auftreten, haben sie auch einen eigenen Namen bekommen. Wir lernen nun einige spezielle Verteilungen kennen.

Spezielle Verteilung 1: Binomialverteilung – „Anzahl Erfolge“

Oft im Leben gibt es *zwei* Möglichkeiten: Bestehen (= **Erfolg**) oder Durchfallen (= **Misserfolg**).

Genau diese Situation wird durch die Bi (= *zwei*) nomialverteilung beschrieben.

Wir merken uns zwei Beispiele:

- **klassische Wahrscheinlichkeit: Würfel**
  Einen Würfel können wir n-mal werfen und nach der Anzahl der geworfenen „6“ fragen. Bei jedem Wurf sind also zwei Ausgänge möglich: „6“ und „nicht 6“. Die „Erfolgswahrscheinlichkeit“ ist bei jedem Wurf $p(6) = \frac{1}{6}$.
- **Statistik: Eigenschaft**
  Wir wählen eine Stichprobe von n Personen und interessieren uns für eine Eigenschaft. Bei jeder Person sind also zwei Ausgänge möglich: die Person „hat die Eigenschaft“ oder die Person „hat nicht die Eigenschaft“.
  Die „Erfolgswahrscheinlichkeit“ (oder: „Trefferwahrscheinlichkeit“), also dass eine Person eine bestimmte Eigenschaft besitzt, wird dabei mit einer langen Versuchsreihe bestimmt. Bei der Eigenschaft „Linkshändigkeit“ schätzt man etwa: $p(L) = 0.12$.

Beispiel 13 wie oft kommt die „6“ in 5 Würfen? / wie viele „Linkshänder“ gibt es in einer Klasse?

a) Die Zufallsgrösse X ist wie folgt definiert: X = Anzahl „6“ in 5 Würfen.
Bestimmen Sie die Wahrscheinlichkeitsverteilung! Skizzieren Sie die Verteilung!

b) Die Zufallsgrösse X sei gegeben durch X = Anzahl „Linkshänder“ in einer Stichprobe von n = 20.
Bestimmen Sie die Wahrscheinlichkeitsverteilung! Skizzieren Sie die Verteilung!

Ein Zufallsversuch hat nur zwei mögliche Ergebnisse („Erfolg“ und „Misserfolg“) und wird mehrmals wiederholt.
Es sei

- n = Anzahl Wiederholungen (z.B. Grösse der Stichprobe)
- p = Erfolgswahrscheinlichkeit (z.B. Person hat Eigenschaft = „Erfolg“ oder „Treffer“)

Wir definieren die Zufallsgrösse

X = Anzahl Erfolge

X nimmt also die Werte 0, 1, 2, ... , n an.

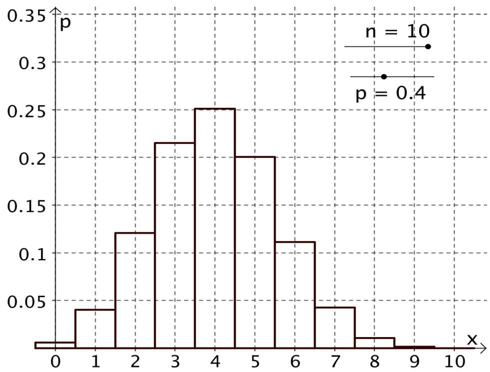
Die Verteilung von X heisst **Binomialverteilung** (mit den Parametern n und p).
Ihre Wahrscheinlichkeiten berechnen sich wie folgt:

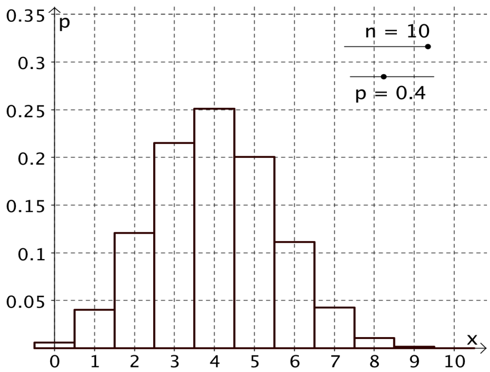
$$\mathbf{p(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}}.$$

Beispiel 14 Binomialverteilung, färben & rechnen, TR

Färben Sie im Histogramm die entsprechende Fläche ein und notieren Sie dann die Wahrscheinlichkeit in *korrekter Schreibweise*. Berechnen Sie sie anschliessend mit Hilfe des TR.

a) Einzelwerte (distr, Binomialpdf)
Binomialverteilung: X = Anzahl Erfolge mit n = 10 und Erfolgswahrscheinlichkeit p = 0.4.

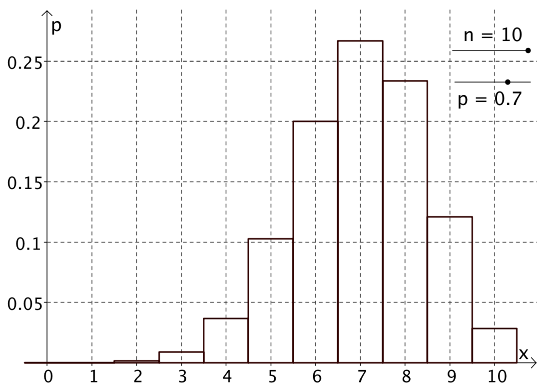


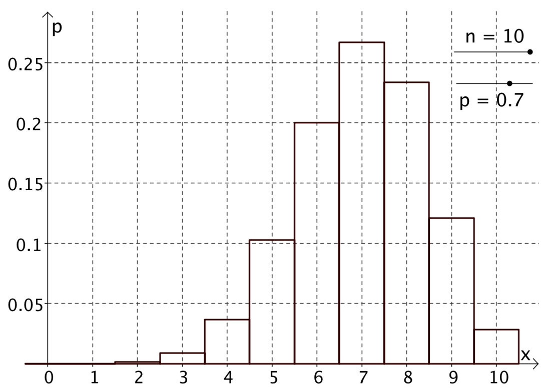


p(X = 3) =

p(X = 7) =

b) summierte Werte (distr, Binomialcdf)
Binomialverteilung: X = Anzahl Treffer mit n = 10 und Trefferwahrscheinlichkeit p = 0.7.





p(X ≤ 8) =

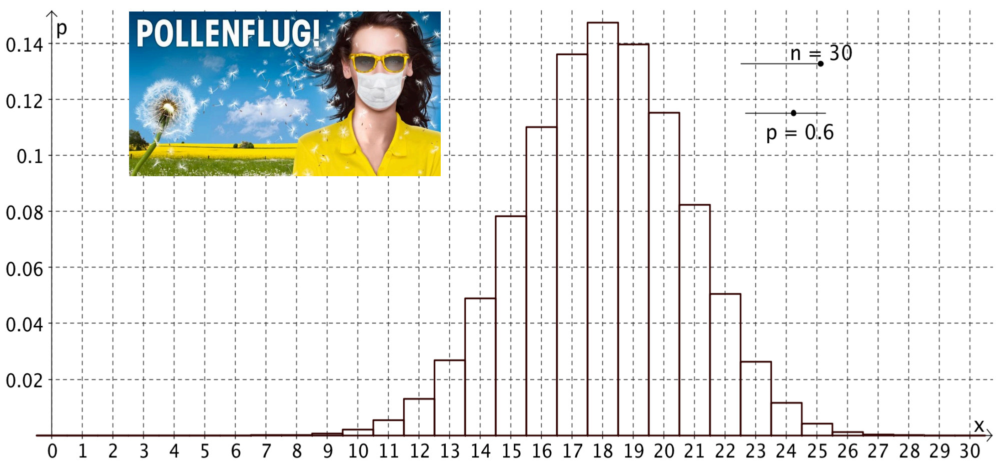
p(5 ≤ X ≤ 8) =

Beispiel 15 Heuschnupfen

a) Färben Sie im Histogramm die entsprechende Fläche ein und notieren Sie dann die Wahrscheinlichkeit in *korrekter Schreibweise*. Berechnen Sie diese anschliessend mit Hilfe des TR.

Pollen können Heuschnupfen auslösen. Ein Nasenspray wirkt in 60% aller Anwendungen lindernd.
30 Patienten (Stichprobe mit n = 30) nehmen das Mittel ein. Wie gross ist die Wahrscheinlichkeit, dass

- *genau* 18 Personen geheilt werden?
- *höchstens* 12 Personen geheilt werden?
- *zwischen* 20 und 24 Personen geheilt werden?



b) Das Histogramm
- hat bei X = 18 die grösste (höchste) Säule. Warum ist das so?
- ist nicht symmetrisch zur Säule bei X = 18. Warum?

Ein Rätsel (?)
Sie werfen eine Münze. Ist es wahrscheinlicher, dass Sie

- in 10 Würfen genau 5-mal „Kopf" werfen oder
- in 100 Würfen genau 50-mal „Kopf" werfen ??

Beispiel 16 qualitative Histogramme / Gedankenexperiment

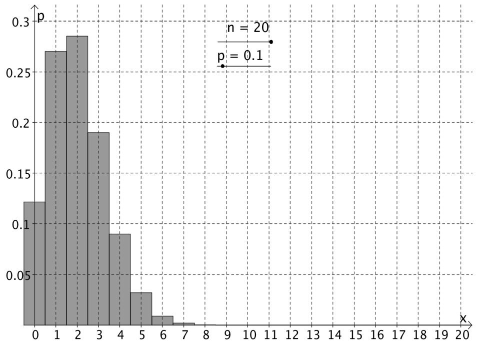
Zeichnen Sie jeweils ein **qualitatives** Histogramm zu folgenden Binomialverteilungen.
Überlegen Sie dazu:
- Wo ist die höchste Säule? Mathematisch formuliert: wo liegt das Maximum der Verteilung?
- Ist das Histogramm (die Verteilung) symmetrisch? Oder: fast symmetrisch?

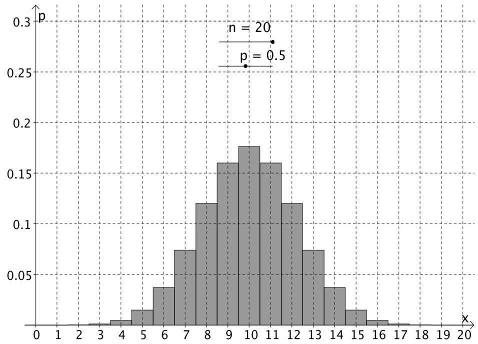
a) Eine Münze wird geworfen. Sei X = Anzahl Kopf.
- Die Münze wird 10-mal geworfen.
- Die Münze wird 100-mal geworfen.
- Die Münze wird 1000-mal geworfen.

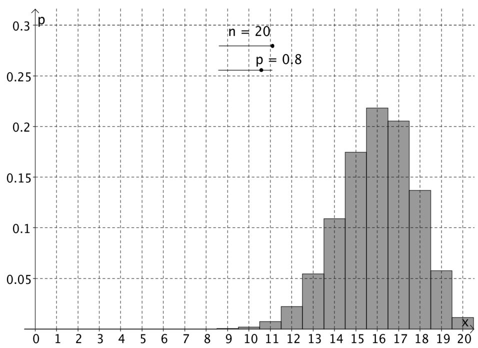
b) Der SVP Wähleranteil beträgt rund 30%. Sei X = Anzahl SVP Wähler in der erhobenen Stichprobe.
- Die Stichprobe umfasst 10 Personen.
- Die Stichprobe umfasst 100 Personen.
- Die Stichprobe umfasst 1000 Personen.

Beispiel 17 Eigenschaften der Binomialverteilung

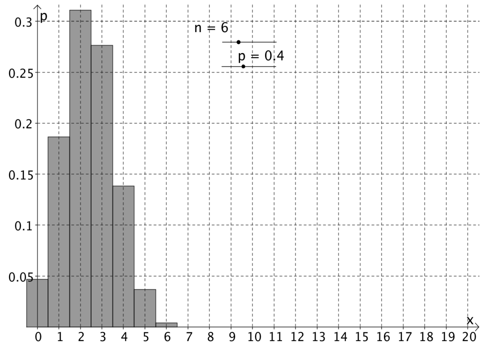
Variieren von p (n fest)

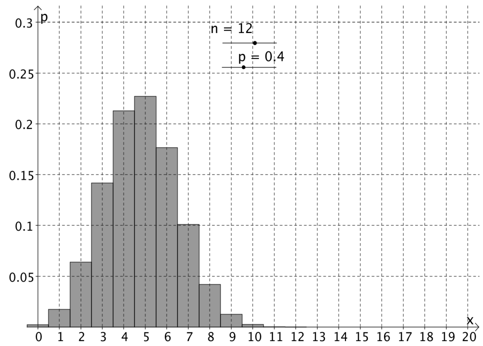


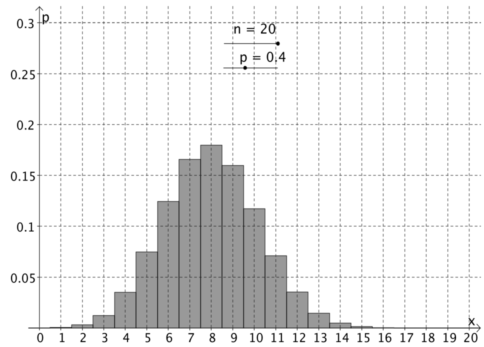




Variieren von n (p fest)







Ergänzen Sie den Lückentext! Die Lösungen sind im Begriffssalat unten angegeben.

- in Abhängigkeit von p (n fest)

  Je grösser p ist, umso weiter ______________________________ liegt das Maximum der Verteilung.

  Für $p = 0.5$ ist die Verteilung ______________________________________

- in Abhängigkeit von n (p fest)

  Je grösser n ist, umso ____ wird die Verteilung.

  Je grösser n ist, umso ____ wird die Verteilung.

Begriffssalat (Lösungen)
symmetrisch; rechts; flacher und breiter; symmetrischer

Erwartungswert und Standardabweichung der Binomialverteilung, Interpretation

Ist X die Zufallsgrösse, die die Anzahl Erfolge bei n Versuchen angibt, so hängen

- der Erwartungswert μ – die durchschnittlich zu erwartende Trefferzahl – und
- die Standardabweichung σ – die „mittlere Abweichung“ vom Erwartungswert –

von den Parametern n und p ab. Dies zeigte sich an den Histogrammen im vorigen Abschnitt.

Die Formeln für den Erwartungswert und die Standardabweichung lassen sich bei binomialverteilten Zufallsgrössen umformen und in eine sehr einfache Form bringen ☺.
Ehrgeizige können versuchen, diese „Formeln“ für allgemeines p und n = 1, 2, 3 zu begründen...!

Erwartungswert: intuitiv!

- Wie lautet die Formel für den Erwartungswert μ einer „allgemeinen“ Zufallsgrösse X?
- Wie lautet wohl die Formel für den Erwartungswert μ einer „binomialverteilten“ Zufallsgrösse X?
- Interpretation: Wo „sieht“ man den Erwartungswert im Histogramm?

Standardabweichung: nicht intuitiv!

- Wie lautet die Formel für die Standardabweichung σ einer „allgemeinen“ Zufallsgrösse X?
- Wie lautet wohl die Formel für die Standardabweichung σ einer „binomialverteilten“ Zufallsgrösse X? (Von welchen Grössen hängt die Formel wohl ab?)
- Interpretation*: Wo „sieht“ man die Standardabweichung im Histogramm?

* **Interpretation der Standardabweichung**

Wir wissen, dass die Histogramme von Binomialverteilungen immer flacher und breiter werden, wenn die Anzahl n der Wiederholungen zunimmt.
Dem Maximum (= dem Erwartungswert μ) einer solchen Verteilung kommt keine allzu grosse Wahrscheinlichkeit mehr zu; vielmehr treten die Nachbarswerte des Maximums mit vergleichbaren Wahrscheinlichkeiten auf.
Da die Histogramme bei grossem Stichprobenumfang n nahezu *symmetrisch* zum Erwartungswert $\mu = n \cdot p$ sind, interessieren wir uns speziell für symmetrische Bereiche um den Erwartungswert μ und deren Wahrscheinlichkeiten. Jetzt kommt die Standardabweichung σ ins Spiel!
Die Standardabweichung σ ist ja ein Mass dafür, wie die Werte um den Erwartungswert streuen.
Überprüfen Sie am folgenden Beispiel, dass „in etwa“ folgendes gilt:

„Quick-and-dirty Regeln“

1σ-Umgebung: $P(\mu - 1\sigma \le X \le \mu + 1\sigma) \approx 2/3$
2σ-Umgebung: $P(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 95\%$
3σ-Umgebung: $P(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx$ fast alle

Beispiel 18 „Quick-and-dirty“ Regeln verifizieren

Berechnen Sie Erwartungswert und Standardabweichung. Verifizieren Sie die obigen Regeln. Interpretieren Sie damit die Standardabweichung.

- Eine Münze wird n = 100 mal geworfen. Sei X = Anzahl „Kopf“.
- Ein Würfel wird n = 200 mal geworfen. Sei X = Anzahl „6“.
- Der SVP Wähleranteil beträgt rund 30%. Sei X = Anzahl SVP Wähler in der Stichprobe. Es wird eine Stichprobe von n = 1000 Personen erhoben.

Später – im Abschnitt über die sog. „Normalverteilung“ – werden diese Regeln präzisiert.
Beachte *diese Werte gelten (für genügend grosse n) praktisch unabhängig von n und p.*

Spezielle Verteilung 2: Poisson-Verteilung – falls nur μ bekannt ist

Simeon Poisson, 1781-1840

Anwendung 1: Approximation der Binomialverteilung

Für grosse n ($n > 100$) und kleine p ($p < 0.1$) fand Poisson eine überraschend gute Approximation für die Binomialverteilung.

Ein Zufallsversuch hat nur zwei mögliche Ergebnisse („Erfolg“, „Misserfolg“) und wird (unendlich) oft wiederholt. Wir definieren die Zufallsgrösse

X = Anzahl Erfolge

X nimmt also die Werte 0, 1, 2, ... an.

Die Verteilung von X heisst **Poisson-Verteilung** (mit den Parameter μ), wenn sich die Wahrscheinlichkeiten für die Werte (Anzahl Treffer) wie folgt berechnen:

$$\mathbf{p(X = x) = \frac{\mu^x}{x!} \cdot e^{-\mu} \qquad (x \in IN)}$$

Beachte *Die Poisson-Verteilung hängt von einem einzigen Parameter – dem „Erwartungswert“ $\mu = n \cdot p$ – ab.*

Beispiel 19 **Farbenblindheit / Geburtstag**

Rechnen Sie einmal mit der Binomialverteilung und einmal mit der Poissonverteilung. Vergleichen Sie.

a) Farbenblindheit tritt in der Bevölkerung mit 1% auf. Mit welcher Wahrscheinlichkeit sind von 200 Personen

- genau 3
- weniger als 3
- mehr als 4 farbenblind.

b) Berechnen Sie die Wahrscheinlichkeit, dass von 365 Personen genau

- eine
- mindestens 3 am gleichen Tag Geburtstag haben wie Sie.

(Nicht zu verwechseln mit dem Geburtstagsparadoxon.)

Da die Poissonverteilung dann zum Zuge kommen kann, wenn p sehr klein ist, heisst sie manchmal auch die **Verteilung der seltenen Ereignisse**.

Ein Rätsel (?)
Im Schnitt hat Esmeralda pro Monat 3-mal Kopfschmerzen.
Wie gross ist die Wahrscheinlichkeit, dass sie im kommenden Monat genau 2-mal Kopfschmerzen hat?

Anwendung 2: Poisson-Verteilung im Alltag

Die Poissonverteilung ist aber mehr als nur eine Approximation der Binomialverteilung für grosses n und kleines p. Die Poisson-Verteilung wird vor allem dort eingesetzt, wo die Häufigkeit eines Ereignisses zB. über eine gewisse *Zeitspanne* betrachtet wird. Wir wissen nämlich, dass die Poisson-Verteilung nur von einem Parameter, nämlich dem Erwartungswert μ, bestimmt wird...

Beispiel/Herleitung

Auf einem Strassenstück ereignet sich gemäss polizeilichen Angaben im Mittel 0.4 Unfälle pro Monat.
Mit welcher Wahrscheinlichkeit ereignet sich auf diesem Strassenstück in einem Monat genau ein Unfall?

Beachte *Diese Aufgabe entspricht der eingangs gestellten Rätselaufgabe!*

Wir versuchen ein *mathematisches Modell* für diese Situation aufzustellen.

Modell 1
Wir nehmen an, dass in einem Monat ein Unfall eintritt mit Wahrscheinlichkeit p = 0.4 und kein Unfall mit der Wahrscheinlichkeit 1 – p = 0.6.
Wir haben also eine Binomialverteilung mit den Parametern:

$$n = 1;\ p = 0.4 \quad (\text{und } \mu = 0.4)$$

Dieses Modell erfüllt zwar die Bedingung, dass sich im Mittel 0.4 Unfälle ereignen, aber es entspricht nicht der Realität, da sich in **einem (n = 1)** Monat ja auch mehr als ein Unfall ereignen kann.
Wir müssen also unser Modell verbessern.

Modell 2
Wir nehmen an, dass in einem Tag ein Unfall eintritt mit der Wahrscheinlichkeit
$p = \frac{\mu}{n} = \frac{0.4}{30} = 0.013$ und kein Unfall mit der Wahrscheinlichkeit 1 – p = 0.987. Für die Anzahl der Unfälle in einem Monat haben wir also eine Binomialverteilung mit:

$$n = 30;\ p = 0.013 \quad (\text{und } \mu = 0.4)$$

Dieses Modell erfüllt immer noch die Bedingung, dass sich im Mittel 0.4 Unfälle pro Monat ereignen, ist aber wesentlich näher an der Realität, weil nun täglich (n = 30) ein Unfall passieren kann. Es ist aber immer noch nicht genau, da es ja auch möglich ist, dass am gleichen Tag mehrere Unfälle passieren.
Also verbessern wir unser Modell weiter.

Modell 3
Wir nehmen an, dass innerhalb einer Sekunde ein Unfall eintritt mit Wahrscheinlichkeit
$p = \frac{\mu}{n} = \frac{0.4}{30\cdot 24\cdot 3600} = 0.000000154$ und kein Unfall mit Wahrscheinlichkeit 1 – p = 0.999999846. Für die Anzahl Unfälle in einem Monat haben wir also eine Binomialverteilung mit:

$$n = 30\cdot 24\cdot 3600 = 2'592'000;\ p = 0.000000154 \quad (\text{und } \mu = 0.4)$$

Dieses Modell erfüllt auch immer noch die Bedingung, dass sich im Mittel 0.4 Unfälle pro Monat ereignen, und es bildet die Realität nun beinahe schon exakt ab, da die Wahrscheinlichkeit, dass mehr als ein Unfall pro Sekunde eintritt, verschwindend kein ist. Für unsere gesuchte Wahrscheinlichkeit können wir deshalb schon mit gutem Gewissen annehmen, dass gilt:

$$p(\text{genau ein Unfall}) = \text{Binom}(X = 1) \approx \text{Poisson}(X = 1).$$

Die Binomialverteilung dürfen wir hier sicher durch die Poissonverteilung annähern, da n sehr gross und p sehr klein ist.

Auch wenn dieses Modell schon beinahe der Realität entspricht, können wir es immer noch verbessern, indem wir den Monat in noch kleinere Zeitintervalle aufteilen.
Dies führt uns zum abschliessenden Modell 4.

„Binomial“: n = Anzahl Versuche
„Poisson“: n = Anzahl „Versuche“ = Anzahl „Zeitspannen“

Modell 4
Wir teilen den Monat in noch mehr Zeitintervalle auf und betrachten den Grenzwert $n \to +\infty$. Dadurch werden die Zeitintervalle, in denen ein Unfall eintreten kann unendlich klein, und wir haben ein exaktes Abbild der Realität. Unsere Binomialverteilung sieht dann wie folgt aus:

$n \to +\infty$; $p = \frac{\mu}{n} \to 0$ ($\mu = np = 0.4$)

Somit gilt für die gesuchte Wahrscheinlichkeit:

p(genau ein Unfall) = Binomial(X = 1) = Poisson(X = 1)

Da für $n \to +\infty$ die Poissonverteilung exakt mit der Binomialverteilung übereinstimmt (vgl. Anhang 3), ist die Zufallsgrösse, welche die Anzahl Unfälle pro Monat „zählt“, in der Tat exakt poissonverteilt mit dem Parameter $\mu = np = 0.4$. Also:

$$\text{p(genau 1 Unfall pro Monat)} = \text{Poisson}(X = 1) = \frac{\mu^1}{1!} \cdot e^{-0.4} = 0.27$$

Natürlich können wir auch berechnen, dass sich zB. mehr als 1 Unfall pro Monat ereignet:

$$\text{p(mehr als 1 Unfall pro Monat)} = \text{poisson}(X > 1) = 1 - p(X \le 1) = \ldots = 1 - 0.94 = 0.06$$

Beispiel 20 Raubüberfall / Pizza / Druckfehler

a) In einer Stadt zählt die Polizei im Mittel pro Jahr fünf Raubüberfälle.
Wie gross ist die Wahrscheinlichkeit, dass im nächsten Jahr

- genau 5
- weniger als 5
- mehr als 10 Raubüberfälle statt finden?

b) Ein Pizza-Schnellimbiss führt eine Statistik über die Anzahl der Gäste, die Bestellungen aufgaben.
Laut dieser Statistik kommen jeweils zwischen 18 Uhr und 19 Uhr im Schnitt 12.1 Bestellungen rein.
Mit welcher Wahrscheinlichkeit beträgt die Anzahl Bestellungen zwischen 18 Uhr und 19 Uhr

- genau 8
- höchstens 10
- zwischen 9 und 15 (beide inklusive)
- mindestens 11

c) Ein 500 seitiges Buch hat im Schnitt ca. 50 Druckfehle (die sich zufällig über die Seiten verteilen).

- Wie gross ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Seite mehr als einen Druckfehler hat?
- Mit wie vielen Seiten mit mehr als einem Druckfehler ist im Schnitt zu rechnen?

Natürlich besitzt – so wie jede Verteilung – auch die Poissonverteilung einen Erwartungswert (welchen wohl?) und eine Standardabweichung. Näheres finden Sie dazu im Anhang 3.

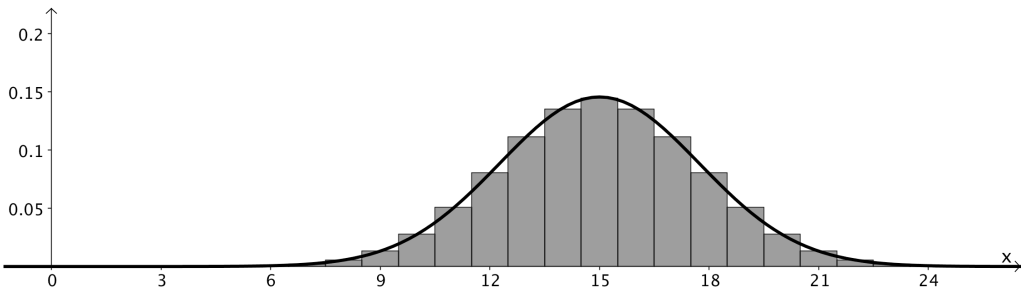
Spezielle Verteilung 3: Normalverteilung – „wichtigste“ Verteilung

de Moivre, 1733; Laplace, 1810

Die Normalverteilung ist die *wichtigste* Verteilung. Dies verrät schon der Name.
Mit ihrer Hilfe lassen sich viele Situationen beschreiben, weil „normal“ (durchschnittlich) im Leben eben viel häufiger vorkommt, als „nicht-normal“.
Viele von uns sind zB. durchschnittlich schön oder intelligent, aber nur wenige sehr oder sehr wenig.
Wir werden sehen...

Anwendung 1: Approximation der Binomialverteilung

Binomialverteilung
n = 30
p = 0.5
z.B. 30-maliges Werfen einer Münze,
X = Anzahl „Kopf“



In das Histogramm einer Binomialverteilung kann eine Kurve gelegt werden, die die Verteilung so gut approximiert, dass die Berechnung von $P(a \le X \le b)$ durch eine Flächenberechnung erfolgen kann.

Diese „Kurve“ heisst **Normalverteilung** oder auch „Gauss’sche Glockenkurve“.

1777 – 1855

Beispiel 21 „Herleitung“ der Normalverteilung (Dichtefunktion)

a) Zuerst brauchen wir eine Funktion, deren Graph eine (symmetrische) „Glockenform“ hat...

b) Welche weiteren „Eigenschaften“ muss die Kurve haben?
Tipp: wie oben angetönt, sollen „Wahrscheinlichkeiten“ mit Hilfe von „Flächen“ berechnet werden...

c) Die Binomialverteilung ist bestimmt durch die Parameter n und p. Die jeweiligen Histogramme werden „charakterisiert“ durch

- die Lage des Maximums (Erwartungswert $\mu = np$) und
- die Breite der Verteilung (Standardabweichung $\sigma = \sqrt{np(1-p)}$)

Soll die Kurve die Binomialverteilung approximieren, muss sie diese beiden Eigenschaften berücksichtigen. D.h.: in der Gleichung treten μ und σ als Parameter auf.

Gleichung der Normalverteilung (Dichtefunktion)

Beispiel 22 Normalverteilungen zeichnen, μ und σ schätzen

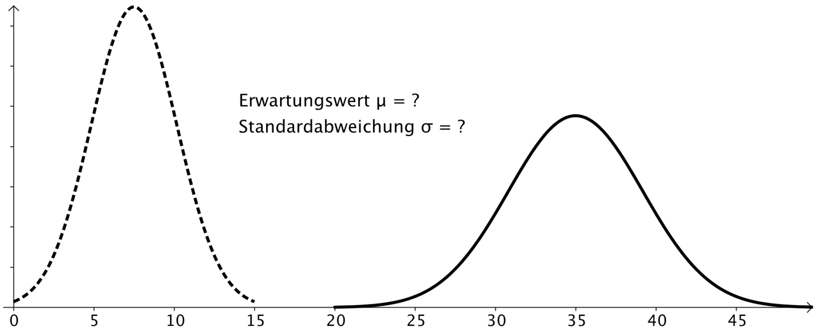
a) Zeichnen Sie (in das gleiche Koordinatensystem) je drei Normalverteilungen,

- mit gleichem Erwartungswert μ, aber unterschiedlicher Standardabweichung σ.
- mit gleicher Standardabweichung σ, aber unterschiedlichem Erwartungswert μ.

b) Man kann zeigen (vgl. Anhang 2):

- die Maximumsstelle der Normalverteilung liegt bei $x = \mu$
- die *Wendestellen* der Normalverteilung liegen bei $x = \mu \pm \sigma$.

Schätzen Sie damit die Parameter μ und σ bei den beiden abgebildeten Normalverteilungen



c) Skizzieren Sie eine Normalverteilung mit $\mu = 100$ und $\sigma = 15$. (Normalverteilung des **IQ**!)

Jetzt wollen wir Wahrscheinlichkeiten mit Hilfe der Normalverteilungen konkret berechnen.

Hinweis *Skizzieren Sie die Verteilung immer qualitativ. Färben Sie gesuchte Wahrscheinlichkeiten ein.*

Beispiel 23 Normalverteilung als Approximation der Binomialverteilung / Stetigkeitskorrektur

Eine Münze wird 50-mal geworfen. Wie gross ist die Wahrscheinlichkeit, dass

a) zwischen 20 und 30 mal „Kopf“ erscheint, also $P(20 \leq X \leq 30)$ und zwar

- „exakt“ mit Hilfe der Binomialverteilung.
- „approximativ“ mit Hilfe der Normalverteilung.

b) Was versteht man unter dem Begriff „**Stetigkeitskorrektur**“?

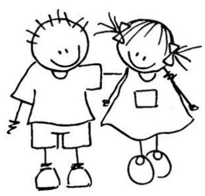
Jede Binomialverteilung (auch für $p \neq 0.5$) wird mit zunehmend grossem n immer symmetrischer und lässt sich durch die (symmetrische) Normalverteilung approximieren. Es lässt sich zeigen:

Die Approximation gilt als hinreichend gut, wenn $\sigma > 3$ ist.

Beispiel 24 Normalverteilung als Approximation der Binomialverteilung

Die statistische Wahrscheinlichkeit für eine Mädchengeburt beträgt 0.486 (!).
Wie wahrscheinlich ist es, dass sich unter den nächsten 1000 Geburten in einem Spital

a) zwischen 450 und 500 Mädchen befinden?
(Ist die Bedingung $\sigma > 3$ erfüllt? Weiter: Stetigkeitskorrektur beachten.)



b) mehr als 500 Mädchen befinden?

Beispiel 25 Normalverteilung als Approximation der Binomialverteilung / invNormal

Ein Medikament hat eine Heilungswahrscheinlichkeit von 80%. 400 Patienten werden behandelt.

a) Wie gross ist die Wahrscheinlichkeit, dass

- höchstens 310 Patienten
- zwischen (einschliesslich) 305 und 335 geheilt werden?

b) In welchem zum Erwartungswert symmetrischen Bereich liegt mit 80% Wahrscheinlichkeit die Anzahl der Geheilten?

Hinweis *wie immer Skizze... und TR-funktion invNormal.*

Das folgende Beispiel hat etwas Schockierendes. Denken Sie darüber nach...

Beispiel 26 Dominanz der Minderheit

Die 200 Mitglieder des Tennis-Clubs wählen einen Präsidenten. Es melden sich nur zwei Bewerber: Michael und Rodscher.
Es handelt sich um eine einfache Mehrheitswahl ohne Möglichkeit der Enthaltung.
Die *beiden* Kandidaten sind eher profillos, sodass die Wahlchancen ausgeglichen erscheinen.
Kurz vor der Wahl aber gewinnt Michael die Clubmeisterschaft.
Das beeindruckt 20 Mitglieder derart, dass diese spontan beschliessen, ihre Stimme Michael zu geben.
Wie verändern sich die Wahlchancen der beiden Kandidaten?

Anwendung 2: stetige Zufallsgrössen

Bisher haben wir nur Zufallsgrössen kennengelernt, die „**zählen**" – nämlich die „Anzahl Erfolge".
Solche Zufallsgrössen heissen **diskret**.

(„diskret" im Sinne von „einzelne", isolierte Werte)

Es gibt aber auch Zufallsgrössen, die „**messen**" – zB. die Körpergrösse oder die Abfüllmenge von Bierflaschen.
Solche Zufallsgrössen heissen **stetig.**

(„stetig" im Sinne von „kontinuierlich")

Stetige Zufallsgrössen können also in einem Intervall *jeden beliebigen reellen Zahlenwert* annehmen.

Stetige Zufallsgrössen sind „von Natur aus" häufig normalverteilt, weil:
Zufallsgrössen, die sich aus einer grossen Anzahl zufallsbedingter, unabhängiger Einflüsse zusammensetzen, sind stets normalverteilt. Das ist einsichtig, wenn man sich die Binomialverteilung vor Augen hält, die sich aus vielen einzelnen Erfolg/Misserfolg-Möglichkeiten zusammensetzt.

Beispiel
Wer ist „gut" in der Schule? „Gut" sein setzt sich zusammen aus Intelligenz, Arbeitseifer, Lehrperson, Schlaf, Geld für Nachhilfestunden etc.
Diese Eigenschaften treten „zufällig" auf. Wenige haben keine oder alle. Viele haben einige. Es ergibt sich eine „Glockenkurve"...

Aus den Messwerten kann man den Erwartungswert und die Standardabweichung „schätzen" und dann mit Hilfe der Normalverteilung verschiedene Problemstellungen lösen.

Beispiel 27 Normalverteilung als stetige Zufallsgrösse

Eine Studie bei 2000 männlichen Studenten ergab ein (geschätztes!) Gewicht von $\mu = 77.5$ kg und eine (geschätzte!) Standardabweichung von $\sigma = 10$ kg.

a) Wie viele Studenten* haben (vermutlich) ein Gewicht zwischen 75 kg und 87.5 kg?

b) Wie viele Studenten wiegen weniger als 60 kg? Wie viele mehr als 100 kg?

c) Die schwersten 2.5% Studenten dürfen bei der Realityshow *„fat, but clever“* mittu(rne)n.
Wie viel muss ein solcher Student auf die Waage bringen?

* Interpretation der Wahrscheinlichkeit:
Die Wahrscheinlichkeit ist als prozentualer Anteil *aller* Studenten zu interpretieren.

Beispiel 28 nur Intervalle sind sinnvoll...

Untersuchungen betreffend Brenndauer einer Glühbirne haben ergeben: $\mu = 80$ h, $\sigma = 5$ h.

a) Wie gross ist die Wahrscheinlichkeit, dass eine Glühbirne
- höchstens 78 h brennt?
- genau (!) 78 h lang brennt? *

b) Bei welcher Brenndauer gehört eine Glühbirne zu den „schlechtesten“ 20%?

* Es zeigt sich hier, dass Wahrscheinlichkeiten über die Normalverteilung nur in *Intervallen* sinnvoll berechnet werden können.
Wie müsste man rechnen, wenn man die Messung „diskretisiert“, also zB. nur die Anzahl *ganzer* Stunden zählt?

Beispiel 29 Bestätigung der sigma-Regeln

a) Die Standardabweichung σ der Binomialverteilung haben wir (vgl. Beispiel 17, „quick-and-dirty Regel“) wie folgt interpretiert:

- Binomial($\mu - 1\sigma \le X \le \mu + 1\sigma$) $\approx 2/3$
- Binomial($\mu - 2\sigma \le X \le \mu + 2\sigma$) $\approx 95\%$
- Binomial($\mu - 3\sigma \le X \le \mu + 3\sigma$) $\approx$ fast alle

Bestätigen Sie diese Regeln mit Hilfe der Normalverteilung. Wie lauten die „genauen“ Werte?
Berechnen Sie also

- Normal($\mu - ...\sigma \le X \le \mu + ...\sigma$) = ...

Sie können dazu μ und σ frei wählen.

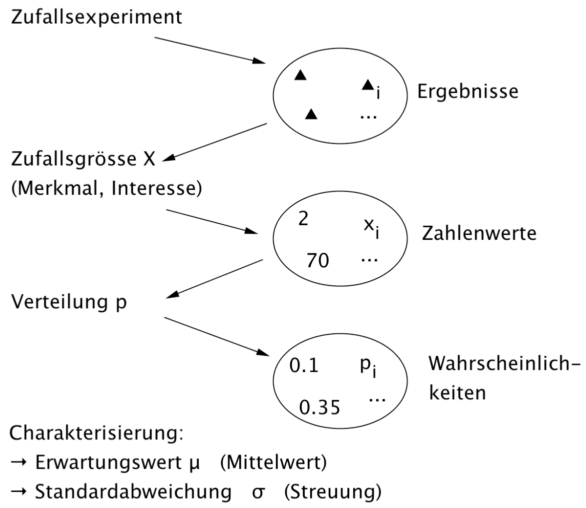
b) Umgekehrt: Bestimmen Sie den Wert z so, dass gilt:

- Normal($\mu - z\cdot\sigma \le X \le \mu + z\cdot\sigma$) = 90%
- Normal($\mu - z\cdot\sigma \le X \le \mu + z\cdot\sigma$) = 95%
- Normal($\mu - z\cdot\sigma \le X \le \mu + z\cdot\sigma$) = 99%

Diese „**z-Werte**“ werden wir später in der Statistik wieder antreffen. Wir merken uns aber schon mal:

$\pm 2\sigma$ (oder eben genauer: $\pm 1.96\sigma$) entspricht 95%.

3 Zusammenfassung



Ein Zufallsversuch „produziert“ Ergebnisse.

Einem solchen Ergebnis kann man eine Zahl zuordnen
(zB. einer zufällig ausgewählten Person ihr Gewicht).

Dieser Zuordnung sagen wir Zufallsgrösse, weil:
Je nach Ergebnis nimmt sie einen anderen Wert an.
Weil das Ergebnis zufällig ist, ist auch der Wert zufällig.

Jeder Wert wird mit einer bestimmten Wahrscheinlichkeit angenommen.
Die Verteilung „verteilt“ die Wahrscheinlichkeiten auf die einzelnen Werte.

Jede Verteilung X wird „charakterisiert“ durch
- den Erwartungswert $\mu(X)$
- die Standardabweichung $\sigma(X)$

Der Erwartungswert entspricht dem Mittelwert.
Die Standardabweichung entspricht der Streuung (= der mittleren Abweichung vom Mittelwert!).

Beachte
Beim Wurf eines Würfels nimmt die Zufallsgrösse die Werte 1, 2, 3, 4, 5 und 6 an.
Jeder Wert ist gleichwahrscheinlich zu p = 1/6 (man spricht auch von einer Gleichverteilung). Weiter ergibt sich
- $\mu(X) = 3.5$
- $\sigma(X) = 1.7$

Was bedeuten nun die Werte $p = 1/6$, $\mu(X) = 3.5$, $\sigma(X) = 1.7$?

Es sind „theoretische“ Werte, die dann zum Tragen kommen, wenn wir ganz oft werfen („auf lange Sicht“).
Wenn wir sehr oft werfen würden, dann würden 1/6 aller Würfe eine bestimmte Augenzahl anzeigen und
- *der Durchschnitt aller Augenzahlen 3.5 ergeben (oder: die zu „erwartende“ Zahl 3.5 betragen)*
- *die geworfenen Augenzahlen im Schnitt um 1.7 von diesem Durchschnitt abweichen*

Es gibt Verteilungen, die sehr häufig auftreten und deshalb einen eigenen Namen tragen.

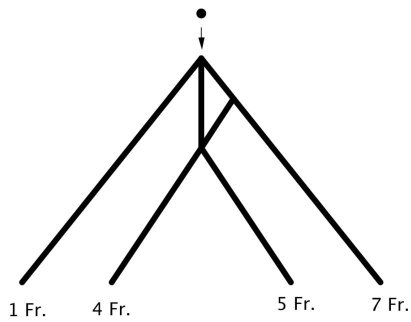
Wir haben kennengelernt:
- Binomialverteilung (abhängig von den Parametern n und p)
- Poisson-Verteilung (abhängig vom Parameter μ)
- Normalverteilung (abhängig von den Parametern μ und σ)

Man unterscheidet zwischen
- diskreten Verteilungen, welche „zählen“ (An*zahl* Erfolge; Binomial- und Poisson-verteilung)
- steigen Verteilungen, welche „messen“ (Normalverteilung)

Die **Normalverteilung** ist die wichtigste Verteilung, weil:
- Häufig ist eine Situation abhängig von vielen verschiedenen Einflüssen, welche sich „im Normalfall“ gegenseitig ausgleichen und damit zur Mitte tendieren.
  Solche Situationen werden von der Normalverteilung modelliert.
- Die Binomialverteilung lässt sich mit Hilfe der Normalverteilung approximieren (Stetigkeitskorrektur!).

Grundaufgabe 1 Wahrscheinlichkeitsverteilung

Eine Kugel rollt im rechts gezeichneten Röhrensystem von oben nach unten.
Bei jeder Verzweigung wählt sie einen der möglichen Wege zufällig.
Je nach Ort, wo die Kugel landet, wird der entsprechende Betrag ausbezahlt.
Sei X = ausbezahlter Betrag.



a) Geben Sie die Wahrscheinlichkeitsverteilung von X an.

b) Berechnen und interpretieren Sie den Erwartungswert und die Standardabweichung von X.

Grundaufgabe 2 Binomialverteilung / Poissonverteilung

a) Ein Medikament hat in 15% aller Fälle eine Nebenwirkung. 30 Patienten nehmen das Medikament.
Wie gross ist die Wahrscheinlichkeit, dass

- genau 2
- zwischen 3 und 7 Patienten Nebenwirkungen zeigen? (3 und 7 inklusive)

b) Ein Multiple-Choice-Test besteht aus 20 Aufgaben, bei denen jeweils nur eine von drei Antworten richtig ist.
Der Test gilt als bestanden, wenn man mehr als die Hälfte der Fragen richtig beantwortet.
Jemand ist schlecht vorbereitet und beantwortet die Fragen durch blosses Raten.
Wie gross ist die Wahrscheinlichkeit, dass dieser Jemand den Test besteht?

c) Wie viele Tore schiesst der FCZ im nächsten Spiel?
Der Statistiker hat den Überblick über die letzten 100 Spiele und hat sich notiert: insgesamt 120 Tore.
Der Herr Canepa möchte wieder einmal zweimal jubeln in einem Spiel. Wie gross ist diese Wahrscheinlichkeit?

Grundaufgabe 3 Normalverteilung

a) Ca. 30% aller Menschen besitzen die Blutgruppe 0. Berechnen Sie die Wahrscheinlichkeit, dass in einer Gruppe von 300 Personen höchstens 100-mal die Blutgruppe 0 auftritt und zwar mit Hilfe

- der Binomialverteilung.
- der Normalverteilung.

b) Die beiden Werte in a) unterscheiden sich praktisch nicht, aber welcher Wert ist eigentlich genauer?

c) Nach einer gängigen Definition gilt ein Haushalt als arm, wenn er über weniger als 50% des Durchschnitts-einkommens verfügt. In der Schweiz beträgt das durchschnittliche Einkommen μ = 7'200 Fr. bei einer Standardabweichung von σ = 3'500 Fr.

- Welcher Anteil der Haushalte gilt demnach als arm?
- Wie hoch ist das Einkommen der Reichsten 5% mindestens?

Hinweis *Wir gehen davon aus, dass das Einkommen normalverteilt ist.*

4 Weitere Aufgaben – Lösungen

Aufgabe 1 IQ-Test

IQ-Tests sind oft so genormt, dass sie einer Normalverteilung folgen mit μ = 100 und σ = 15.
Die Bedeutungen sind in der Tabelle rechts angegeben:
Wie gross ist die Wahrscheinlichkeit, dass

IQ =	55	extrem niedrig
	70	sehr niedrig
	85	niedrig
	100	durchschnittlich
	115	hoch
	130	sehr hoch
	145	extrem hoch

a) ein beliebiger Absolvent mindestens 130 erreicht?

b) in einer Gruppe von 20 getesteten Personen mindestens eine mindestens 130 Punkt erreicht?

Aufgabe 2 Abstimmung

Eine Volksabstimmung soll mit einfacher Mehrheit über eine Gesetzesänderung entscheiden, der die rund 4 Millionen Stimmberechtigten gleichgültig gegenüberstehen. Allerdings ist eine relativ kleine Interessen-gruppe von ca. 3000 Personen wild entschlossen, gegen die Gesetzesänderung zu stimmen.
Mit welcher Wahrscheinlichkeit setzt sie ihren Willen durch?

Aufgabe 3 3-facher Münzwurf

Drei Münzen werden zusammen geworfen. Das Interesse gilt der Wurfkombination „ZZZ“.

a) Wie gross ist die Wahrscheinlichkeit, dass ZZZ im dritten Wurf zum ersten Mal erscheint?

b*) Wie oft muss man im Schnitt warten, bis diese Kombination zum ersten Mal erscheint?

Aufgabe 4 Theorie

a) Wie wirkt sich die Standardabweichung auf das Aussehen der Normalverteilung(-skurve) aus?

b) Was ist der Unterschied zwischen einer diskreten und einer stetigen Verteilung?

c) Jemand kennt die Begriffe: relative Häufigkeit, Mittelwert und Streuung.
Jetzt kommen Sie und erzählen von: Wahrscheinlichkeit, Erwartungswert, Standardabweichung...
Erklären Sie dem jemand, welche Zusammenhänge zwischen diesen „Begriffspaaren“ bestehen.

d) In einem Statistikformelbuch lesen Sie:

$$\mu = \sum x_i p_i \quad \text{und} \quad \sigma = \sqrt{\sum (x_i - \mu)^2 \cdot p_i}\,.$$

Erläutern Sie den *Aufbau* dieser Formeln.

e) Für die Binomialverteilung gilt: $\mu = np$ und $\sigma = \sqrt{np(1-p)}$. Leiten Sie diese Formeln her für allgemeines p und

- n = 1
- n = 2
- evt. n = 3

Lösungen

Aufgabe 1

a) 0.023

b) p(mindestens 1 von 20) = 1 – p(keiner von 20) = $1 - (1 - 0.023)^{20} = 0.37$

Aufgabe 2

93.3% (Approximation der Binomialverteilung durch Normalverteilung)

Aufgabe 3

Geometrische Verteilung (Parameter $p = \frac{1}{8}$)

a) $p(X = 3) = \left(\frac{7}{8}\right)^2 \cdot \frac{1}{8} = 0.096$

b) $\mu = \frac{1}{1/8} = 8$

Aufgabe 4

a) Je grösser σ, umso breiter die Kurve (und entsprechend kleinerem Maximum) und umgekehrt.

b)

- diskret: zählen, bloss isoliert Werte (meistens eine Teilmenge der natürlichen zahlen IN)
  Wahrscheinlichkeiten werden über Summen berechnet
  Bsp: Binomialverteilung
- stetig: messen, Werte aus einem Intervall (Teilmenge der reellen zahlen IR)
  Wahrscheinlichkeiten werden über Integrale berechnet
  Bsp: Normalverteilung

c) ...

d) ...

e) ...

Anhang 1 Geometrische Verteilung

Diese Verteilung ist einfach ☺. Lesen Sie, lösen Sie!

Erinnerung
Sie kennen die **Binomialverteilung** zu den Parametern n und p. Die Zufallsgrösse X ist dabei definiert als

X = Anzahl Erfolge bei n Versuchen

Die Verteilung ist gegeben durch

$$p(X = x) = \binom{n}{x} p^x(1 - p)^{n-x} \quad ,$$

wobei x = Anzahl Erfolge in n Versuchen und p die Erfolgswahrscheinlichkeit bei jedem Versuch ist.

Konkretes Beispiel
Ein Schütze trifft ein Ziel mit der „Erfolgs“wahrscheinlichkeit p = 0,3.
Er schiesst 10-mal. Wie gross ist die Wahrscheinlichkeit, dass er genau 4-mal trifft?

$$p(X = 4) = \binom{10}{4} 0.3^4(1 - 0.3)^{10-4} = \binom{10}{4} 0.3^4 \cdot 0.7^6 = 0.20$$

Eine weitere Verteilung, die in engem Zusammenhang steht mit der Binomialverteilung, ist die geometrische Verteilung. Dazu definieren wir folgende Zufallsgrösse

X = Anzahl Versuche bis zum <u>ersten</u> Erfolg

Beachten Sie den Unterschied zur Binomialverteilung!
Als konkretes Beispiel stellen wir uns zB. einen Schützen vor, der sein Ziel mit Wahrscheinlichkeit p trifft.

Beispiel 30 geometrische Verteilung, allgemein

a) Welche Werte kann X annehmen? Geben Sie die Verteilung von X an!

Also: p(X = x) =

b) Von welchem Parameter / welchen Parametern hängt die geometrische Verteilung ab?

Beispiel 31 geometrische Verteilung, konkret

Sie werfen einen Würfel... bis zum ersten Mal eine „6“ kommt. Es liegt also die Zufallsgrösse vor:

X = Anzahl Würfe bis zur ersten „6“

a) Berechnen und interpretieren Sie $p(X = 4)$ und $p(X \leq 2)$.

b) Wie lange müssen Sie wohl „im Schnitt“ warten bis die erste Sechs fällt?
Mit anderen Worten: wie gross schätzen Sie den Erwartungswert $\mu = E(X)$? Können Sie ihn berechnen?

Natürlich besitzt – so wie jede Verteilung – auch die geometrische Verteilung einen Erwartungswert und eine Standardabweichung.

Erwartungswert

Welches Problem tritt auf, wenn wir versuchen, den Erwartungswert einer geometrisch verteilten Zufallsgrösse zu berechnen?

Im folgenden sind zwei Begründungen angegeben, die ohne eine unendliche Summe auskommen.
Lesen Sie beide durch und verstehen Sie beide!
Merken Sie dann eine, sodass Sie den Erwartungswert herleiten können.

Begründung 1

- Der Erwartungswert E(X) lässt sich per Fallunterscheidung zerlegen. Mit Wahrscheinlichkeit p geht die erste Durchführung erfolgreich aus, mit Wahrscheinlichkeit 1 – p erfolglos, aber der Erwartungswert für die Anzahl der dann noch folgenden Durchführungen ist wegen der *Gedächtnislosigkeit* wiederum E(X).
  Also gilt:

  $E(X) = 1p + (1 - p)\cdot(1 + E(X)) = 1 + (1 - p)\cdot E(X)$, also $E(X) = \frac{1}{p}$.

Begründung 2

- Führt man n Experimente durch, so ist der Erwartungswert für die Anzahl erfolgreicher Durchführungen $n\cdot p$. Daher ist der zu erwartende Abstand zwischen zwei erfolgreichen Durchführungen (einschliesslich einer erfolgreichen Durchführung) $\frac{n}{n\cdot p}$, also $E(X) = \frac{1}{p}$.

Standardabweichung

Wir geben die Formel für die Standardabweichung ohne Herleitung an. Sie lautet:

$$\sigma = \sqrt{\frac{1-p}{p^2}}\,.$$

Beachte *die Verteilung beinhaltet als einzigen Parameter die „Erfolgswahrscheinlichkeit“ p.*
Es ist deshalb klar, dass sowohl der Erwartungswert μ als auch die Standardabweichung σ ausschliesslich von p abhängen.

Machen Sie ein Beispiel zur geometrischen Verteilung!

Dazu gehört die Angabe von:
Zufallsgrösse, Verteilung, Erwartungswert, Standardabweichung.

Stetige Schwester

Die geometrische Verteilung besitzt eine „stetige“ Schwester: die sogenannte **Exponentialverteilung**.
Sie dient als Modell, wenn es zB. um Warte*zeiten* geht.
Mehr zu dieser Verteilung finden Sie – natürlich – im Internet.

Anhang 2 Näheres zur Normalverteilung

Extremstelle = μ, Wendestellen = $\mu \pm 1\sigma$

Zeigen Sie durch Ableiten (Kettenregel!), dass für die Normalverteilung $\varphi(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ gilt:

- $\varphi'(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma^3} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot (x-\mu)$
- $\varphi''(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma^5} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot (\sigma^2 - (x-\mu)^2)$

Zeigen Sie dann, dass gilt:

- $\varphi(x)$ besitzt die Extremstelle (offensichtlich) $x = \mu$
- $\varphi(x)$ besitzt die Wendestellen $x = \mu \pm \sigma$

Standardnormalverteilung

Als Standardnormalverteilung bezeichnet man die Normalverteilung mit dem Erwartungswert $\mu = 0$ und der Standardabweichung $\sigma = 1$. Sie ist demnach gegeben durch:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} .$$

Sie ist gewissermassen die „Mutterfunktion“ aller Normalverteilungen.

Die Variable z wird gewählt, um die Standardnormalverteilung von anderen Normalverteilungen abzugrenzen. Die Transformationsgleichung lautet:

$$z = \frac{x-\mu}{\sigma}$$

Mit Hilfe der Standardnormalverteilung lässt sich etwa auch eine Aufgabe wie die folgende lösen:

Ermitteln von σ bei vorgegebener Wahrscheinlichkeit

Mit einer Maschine wird Parfüm in Fläschchen abgefüllt (die Variation der Füllmenge F sei normalverteilt).
Die Maschine ist im Mittel auf $\mu = 100$ ml eingestellt. Wie gross darf bei dieser Produktion die Standardabweichung σ von F höchstens sein, wenn (höchstens) 95% der produzierten Fläschchen im Inhalt höchstens 4 ml vom Erwartungswert μ von F abweichen sollen?

Lösung *$\sigma = 2$ ml*

Anhang 3 Näheres zur Poissonverteilung

Der Name „Poisson" kommt von Simeon Denis Poisson, der sich 1837 mit diese Verteilung beschäftigte.
Den Titel „Verteilung der seltenen Ereignisse" hat sie aufgrund der Idee, die hinter ihr steckt:
Die Poissonverteilung soll die Häufigkeit des Auftretens von Ereignissen beschreiben, die bei einem einzelnen Element sehr selten auftreten. Da aber eine sehr große Anzahl von Elementen existiert, bei der das Ereignis eintreten könnte, ist das Ereignis aber derart beobachtbar, dass ein Wert für das durchschnittliche Auftreten in einem Zeit- oder Raumintervall angegeben werden kann.
Zum Beispiel ist die Wahrscheinlichkeit, dass ein bestimmter Einwohner einer Stadt morgens zwischen 10:00 Uhr und 10:05 die Postfiliale der Stadt betritt, sehr gering (also p klein). Da aber in der Stadt sehr viele Menschen leben (also n gross), liegt die Zahl der Leute, die die Postfiliale betreten, in einer recht anschaulichen und mit unserem Zahlverständnis begreifbaren Größenordnung.
Mathematisch gesehen wird die Poissonverteilung aus der Binomialverteilung hergeleitet.

Herleitung der Poissonverteilung aus der Binomialverteilung

Approximation der Binomialverteilung durch die Poissonverteilung bzw.:

$$\text{für festes } \mu \text{ und } n \to +\infty \text{ ist: } \quad \text{binomial} = \text{poisson}$$

Wir müssen also zeigen:

$$\text{binomial}(X = x) = \binom{n}{x} p^x(1-p)^{n-x} = \ldots = \frac{\mu^x}{x!} \cdot e^{-\mu} = \text{poisson}(X = x)$$

Hinweis *es ist* $p = \frac{\mu}{n}$.

$$\text{binomial}(X = x) = \binom{n}{x} p^x(1-p)^{n-x} = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-x+1)}{x!} \cdot p^x(1-p)^{n-x}$$

$$= \frac{n\cdot(n-1)\cdot\ldots\cdot(n-x+1)}{x!} \cdot \left(\frac{\mu}{n}\right)^x (1 - \frac{\mu}{n})^{n-x}$$

$$= \frac{\mu^x}{x!} \cdot \frac{n}{n} \cdot \frac{n-1}{n} \cdot \ldots \cdot \frac{n-x+1}{n} \cdot (1 - \frac{\mu}{n})^n \cdot (1 - \frac{\mu}{n})^{-x}$$

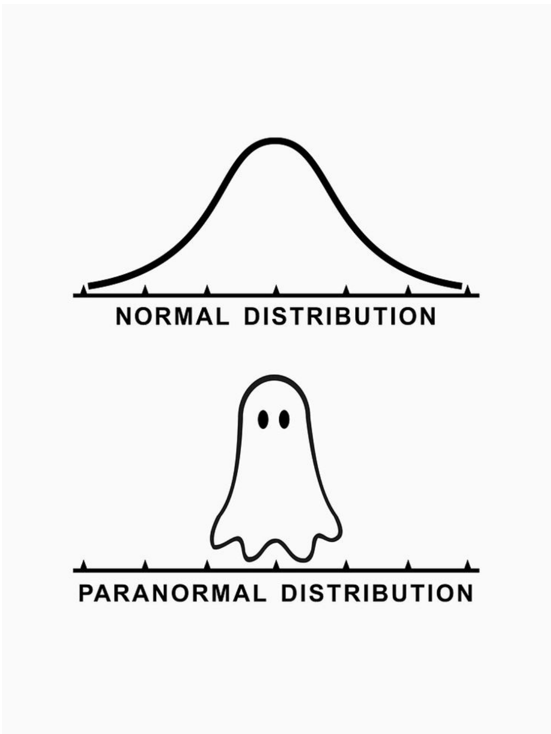
Wir sind nun schon nahe dran! Man kann zeigen, dass gilt: $(1 + \frac{r}{n})^n = e^r$, falls $n \to +\infty$.

Führen Sie die Begründung zu Ende!

Erwartungswert und Standardabweichung der Poissonverteilung

Beachte *die Verteilung beinhaltet als einzigen Parameter μ. Es ist deshalb klar, dass sowohl der Erwartungswert μ als auch die Standardabweichung σ ausschliesslich von μ abhängen.*

- $E(X) = \mu$ (dies ist nicht weiter erstaunlich...)
- $\sigma(X) = \sqrt{\mu}$


NORMAL DISTRIBUTION
PARANORMAL DISTRIBUTION