Normalverteilung

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

Gauß

Normalverteilung

Satz: f aus (1) ist Dichte.

Beweis: 1. $f(x) \geq 0 \; \forall x \in \mathbf{R}$ und $\sigma > 0$.

2. bleibt z.z.

$$\lim_{x \longrightarrow \infty} F(x) = \int_{-\infty}^{\infty} f(t)\, dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt = 1.$$

Wir bezeichnen

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx =: I.$$

Normalverteilung

Wir betrachten zunächst:

$$
\begin{aligned}
I^2 &= \left(\frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx \right)^2 \\
&= \frac{1}{2\pi\sigma^2} \left(\int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx \right) \left(\int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \, dy \right) \\
&= \frac{1}{2\pi\sigma^2} \int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx \right) e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \, dy \\
&= \frac{1}{2\pi\sigma^2} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \, dx \, dy
\end{aligned}
$$

Substitution:

$$s := \frac{x-\mu}{\sigma} \qquad t := \frac{y-\mu}{\sigma}.$$

$$dx = \sigma\, ds \qquad dy = \sigma\, dt.$$

Wir erhalten damit:

$$\begin{aligned} I^2 &= \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}s^2} e^{-\frac{1}{2}t^2} \sigma^2\, ds\, dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(s^2+t^2)}\, ds\, dt \end{aligned}$$

Normalverteilung

Weitere Substitution (Polarkoordinaten):

$$s = r\cos\varphi \qquad t = r\sin\varphi.$$

Dann gilt allgemein nach der Substitutionsregel:

$$\int\int g(s,t)\,ds\,dt = \int\int g(r,\varphi)\det J\,dr\,d\varphi,$$

wobei hier:

$$\begin{aligned}
\det J = |J| &= \begin{vmatrix} \frac{\partial s}{\partial r} & \frac{\partial s}{\partial \varphi} \\ \frac{\partial t}{\partial r} & \frac{\partial t}{\partial \varphi} \end{vmatrix} \\
&= \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} \\
&= r\cos^2\varphi + r\sin^2\varphi \\
&= r(\cos^2\varphi + \sin^2\varphi) = r
\end{aligned}$$

Normalverteilung

$$
\begin{aligned}
I^2 &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}(r^2\cos^2\varphi + r^2\sin^2\varphi)}\, r\,dr\,d\varphi \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}r^2}\, r\,dr\,d\varphi \\
&= \frac{1}{2\pi}\int_0^{2\pi} \left[-e^{-\frac{r^2}{2}}\right]_0^{\infty} d\varphi \\
&= \frac{1}{2\pi}\int_0^{2\pi} d\varphi = \frac{1}{2\pi}2\pi = 1
\end{aligned}
$$

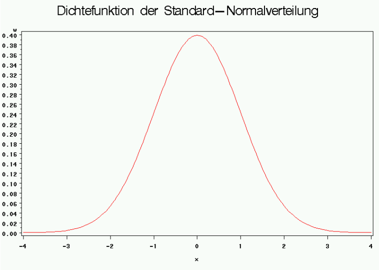
$$\mu = 0, \quad \sigma^2 = 1$$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2} \qquad \text{Dichte}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \qquad \text{Verteilungsfunktion}$$

$\varphi(x), \Phi(x)$ sind tabelliert.
Es geht auch einfacher mit CDF und PDF.

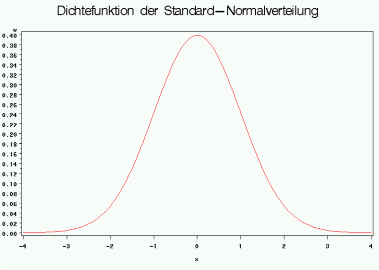
Dichte der Standardnormalverteilung



$$\begin{aligned}\varphi(x) &= \varphi(-x)\\ \Phi(x) &= 1-\Phi(-x)\end{aligned}$$

`Programm: Descr_normal.sas`

Dichte der Standardnormalverteilung



$$\varphi(x) = \varphi(-x)$$
$$\Phi(x) = 1 - \Phi(-x)$$

`Programm: Descr_normal.sas`

Frage: Für welches x gilt: $\Phi(x) = \alpha$?

$x = \Phi^{-1}(\alpha)$ α-Quantil.

$\Phi^{-1}(\alpha)$ als Funktion: Quantilfunktion

SAS: QUANTILE('normal',α,0,1)

Sei $X \sim N(0,1)$. Dann $P(a < X < b) = \Phi(b) - \Phi(a)$.

Satz. Es gilt:

$$
\begin{aligned}
X \sim N(0,1) &\iff \sigma X + \mu \sim N(\mu, \sigma^2) \\
X \sim N(\mu, \sigma^2) &\iff \alpha X + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2) \\
X \sim N(\mu, \sigma^2) &\iff \frac{X - \mu}{\sigma} \sim N(0,1)
\end{aligned}
$$

Normalverteilung

Beziehung zur Standard-Normalverteilung

Sei $X \sim N(0,1)$. Dann $P(a < X < b) = \Phi(b) - \Phi(a)$.

Satz. Es gilt:

$$\begin{aligned} X \sim N(0,1) &\iff \sigma X + \mu \sim N(\mu, \sigma^2) \\ X \sim N(\mu, \sigma^2) &\iff \alpha X + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2) \\ X \sim N(\mu, \sigma^2) &\iff \frac{X-\mu}{\sigma} \sim N(0,1) \end{aligned}$$

Beweis: Wir zeigen nur 1. ($\rightarrow$). Sei $X \sim N(0,1)$.

$$P(\sigma X + \mu \le x) = P\left(X \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right) =$$

$$= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(u-\mu)^2/(2\sigma^2)}\, du$$

Normalverteilung

Unterschiedliche Parameter (1)

Vergleichen Sie

a) σ^2 fest, μ verschieden
b) μ fest, σ^2 verschieden

```
Descr_Normal_1.sas
```

Normalverteilung

Unterschiedliche Parameter (1)

Satz:

Seien $X_1 \sim N(\mu, \sigma_1^2), X_2 \sim N(\mu, \sigma_2^2)$, $\sigma_1^2 < \sigma_2^2$ und $a > 0$. Dann gilt:

$$P(\mu - a < X_1 < \mu + a) > P(\mu - a < X_2 < \mu + a).$$

Beweis:

$$\begin{aligned} P(\mu - a < X_1 < \mu + a) &= P(\frac{-a}{\sigma_1} < \frac{X_1 - \mu}{\sigma_1} < \frac{a}{\sigma_1}) \\ &= \Phi(\frac{a}{\sigma_1}) - \Phi(-\frac{a}{\sigma_1}) \\ &> \Phi(\frac{a}{\sigma_2}) - \Phi(-\frac{a}{\sigma_2}) \\ &= P(\mu - a < X_2 < \mu + a). \end{aligned}$$

Normalverteilung

Beispiel: $X_1 \sim N(10,4), X_2 \sim N(10,9), a = 1.$

$$\begin{aligned}
P(9 < X_1 < 11) &= \Phi(\frac{11-10}{2}) - \Phi(\frac{9-10}{2}) \\
&= \Phi(\frac{1}{2}) - \Phi(-\frac{1}{2}) \\
&= 2 \cdot \Phi(\frac{1}{2}) - 1 \\
&= 2 \cdot 0.6915 - 1 = 0.383.
\end{aligned}$$

$$\begin{aligned}
P(9 < X_2 < 11) &= \Phi(\frac{11-10}{3}) - \Phi(\frac{9-10}{3}) \\
&= \Phi(\frac{1}{3}) - \Phi(-\frac{1}{3}) \\
&= 2 \cdot \Phi(\frac{1}{3}) - 1 \\
&= 2 \cdot 0.6306 - 1 = 0.26112.
\end{aligned}$$

Diskrete Verteilungen

Binomial $X \sim B(n, p)$

X : Anzahl von “Erfolgen”, n Versuche, Erfolgswkt. p.

Poisson $X \sim Poi(\lambda)$

X : Anzahl von “Erfolgen”, n Versuche, Erfolgswkt. p,
n groß und p klein, $n \cdot p = \lambda$.
X : $\#$ Ankünfte in einem Zeitintervall.

Geometrisch, $X \sim Geo(p)$

X :: Zahl der Versuche bis zum ersten “Erfolg”.

Stetige Verteilungen

Gleichverteilung $X \sim R(a, b)$

Zufallszahlen

Exponential $X \sim Exp(\lambda)$

“gedächtnislose” stetige Verteilung.

Normal $X \sim N(\mu, \sigma^2)$

Zentraler Grenzwertsatz
Fehlergesetz (viele kleine unabh. Fehler)

Erwartungswert
Einleitende Motivation

Eine Münze wird 3 mal geworfen.
Wie oft können wir erwarten, daß Blatt oben liegt?
Wie oft wird im Mittel Blatt oben liegen?

$$X: \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1/8 & 3/8 & 3/8 & 1/8 \end{pmatrix}$$

Erwartungswert:
$0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1.5$
D.h. bei 10maliger Durchführung des Experiments können wir im Mittel mit 15mal Blatt rechnen!

Erwartungswert

Diskrete Zufallsvariable

Sei X diskrete Zufallsvariable

$$X : \begin{pmatrix} x_1 & \ldots & x_n & \ldots \\ p_1 & \ldots & p_n & \ldots \end{pmatrix}$$

$$\mathbf{E}X = \sum_{i=1}^{\infty} p_i x_i$$

heißt <u>Erwartungswert</u> von X.

Erwartungswert

$X \sim$ Poisson(λ)

$$X: \begin{pmatrix} 0 & 1 & 2 & 3 & \dots \\ p_0 & p_1 & p_2 & p_3 & \dots \end{pmatrix} \qquad p_i = \frac{\lambda^i}{i!} e^{-\lambda}$$

$$\begin{aligned} \mathbf{E}X &= \sum_{i=0}^{\infty} p_i i \\ &= \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} \cdot i \\ &= \lambda \underbrace{\sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!}}_{e^{\lambda}} e^{-\lambda} = \lambda. \end{aligned}$$

z.B. mittlere Ankunftsrate.

Erwartungswert

$X \sim Bi(n,p)$

$$
\begin{aligned}
\mathbf{E}X &= \sum_{k=0}^{n} k \binom{n}{k} p^k \cdot (1-p)^{n-k} \\
&= p \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} \\
&= p \cdot n \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\
&= p \cdot n \sum_{i=0}^{n-1} \binom{n-1}{i} p^{i} (1-p)^{n-1-i}, \qquad k = i+1 \\
&= n \cdot p.
\end{aligned}
$$

Sei X stetig mit Dichte f. Die Größe

$$\mathbf{E}X = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

heißt Erwartungswert von X.

$X \sim Exp(\lambda), \qquad \lambda > 0$

$$\mathbf{E}X = \int_0^{\infty} x \cdot \frac{1}{\lambda} \cdot e^{-\frac{x}{\lambda}} dx = \lambda$$

Erwartungswert
Normalverteilung

$X \sim N(\mu, \sigma^2)$

$$
\begin{aligned}
\mathbf{E}X &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-(\frac{x-\mu}{\sigma})^2/2} dx \\
&= \int_{-\infty}^{\infty} (\sigma t + \mu) \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} dt \\
&= \mu + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma \cdot t \cdot e^{\frac{-t^2}{2}} dt \\
&= \mu.
\end{aligned}
$$

$$\frac{x-\mu}{\sigma} = t, \quad dx = \sigma dt$$

Erwartungswert
Gleichverteilung

$X \sim R(a,b)$, gleichverteilt auf dem Intervall (a,b)

$$\begin{aligned}
\mathbf{E}X &= \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b \\
&= \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.
\end{aligned}$$

$\mathbf{E}$ ist Linearer Operator

$$\mathbf{E}(aX + bY) = a\mathbf{E}X + b\mathbf{E}Y.$$

Erwartungswert
Eigenschaften des Erwartungswertes

$\mathbf{E}$ ist Linearer Operator

$$\mathbf{E}(aX + bY) = a\mathbf{E}X + b\mathbf{E}Y.$$

Seien X und Y stochastisch unabhängig. Dann

$$\mathbf{E}(X \cdot Y) = \mathbf{E}X \cdot \mathbf{E}Y.$$

$\mathbf{E}$ ist Linearer Operator

$$\mathbf{E}(aX + bY) = a\mathbf{E}X + b\mathbf{E}Y.$$

Seien X und Y stochastisch unabhängig. Dann

$$\mathbf{E}(X \cdot Y) = \mathbf{E}X \cdot \mathbf{E}Y.$$

Regel des Faulen Statistikers

Sei X Zufallsvariable, g: $R \longrightarrow R$ (rechtsseitig) stetig $\Rightarrow$

$$\mathbf{E}(g(X)) = \begin{cases} \sum_{i=0}^{\infty} g(x_i)p_i & \text{, falls X diskret} \\ \int\limits_{-\infty}^{\infty} g(x)f(x)dx & \text{, falls X stetig,} \end{cases}$$

vorausgesetzt die Erwartungswerte existieren.