Mathematik für Biologen

Prof. Dr. Rüdiger W. Braun

Heinrich-Heine Universität Düsseldorf

25. November 2009

Umrechnung auf Standardnormalverteilung

Die Zufallsvariable X sei $N(\mu, \sigma^2)$-verteilt. Dann gelten für $a < b$

$$P(a < X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(a < X) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

Normalapproximation

- Die Zufallsvariable X sei $B(n,p)$-verteilt mit

$$n \cdot p \cdot (1-p) > 9$$

- Dann gilt näherungsweise für natürliche Zahlen $a < b$

$$P(a \leq X \leq b) \cong \Phi\left(\frac{b+\frac{1}{2}-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right) - \Phi\left(\frac{a-\frac{1}{2}-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right)$$

- Wenn $a = 0$ oder $b = n$ ist, braucht man nur einen Term

$$P(a \leq X) \cong 1 - \Phi\left(\frac{a-\frac{1}{2}-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right)$$

$$P(X \leq b) \cong \Phi\left(\frac{b+\frac{1}{2}-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right)$$

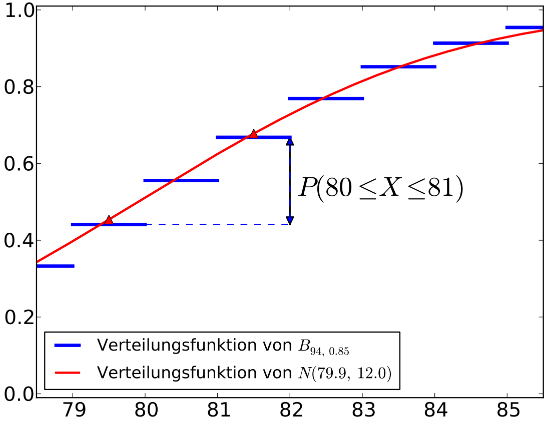
Stetigkeitskorrektur

- Die Normalapproximation ist eine Folge des Grenzübergangs, mit dessen Hilfe ich die Standardnormalverteilung eingeführt hatte.
- Die Terme $\frac{1}{2}$ in den Formeln bezeichnet man als *Stetigkeitskorrektur*. Wir kommen darauf zurück.

Beispiel zur Normalapproximation

- Heilversuch mit 94 Fischen; Heilerfolgswahrscheinlichkeit im Einzelfall $p = 0.85$
- Mit welcher Wahrscheinlichkeit werden mindestens 80 Fische geheilt?
- Die Anwendung der Normalapproximation ist gerechtfertigt, denn
$$n \cdot p \cdot (1 - p) = 94 \cdot 0.85 \cdot 0.15 = 11.985$$
- Also
$$\begin{aligned} P(X \geq 80) &\cong 1 - \Phi\left(\frac{80 - 0.5 - 94 \cdot 0.85}{\sqrt{11.985}}\right) \\ &= 1 - \Phi(-0.1155) = \Phi(0.1155) \cong \Phi(0.12) = 0.548 \end{aligned}$$
- Die Wahrscheinlichkeit, dass mindestens 80 Fische geheilt werden, beträgt ungefähr 55%

Stetigkeitskorrektur: grafische Erklärung



Gleichverteilung

- Eine Zufallsvariable X ist *gleichverteilt* auf dem Intervall $[c, d]$, wenn für je zwei Zahlen $a < b$ aus dem Intervall $[c, d]$ gilt

  $$P(a < X \leq b) = \frac{b - a}{d - c}$$

  Gleichverteilung heißt also, dass die Wahrscheinlichkeit nur von der Länge des Intervalls abhängt.

- Erwartungswert und Varianz für gleichverteiltes X

  $$E(X) = \frac{c + d}{2}, \qquad Var(X) = \frac{(d - c)^2}{12}$$

Dichte und Verteilungsfunktion der Gleichverteilung auf $[c, d]$

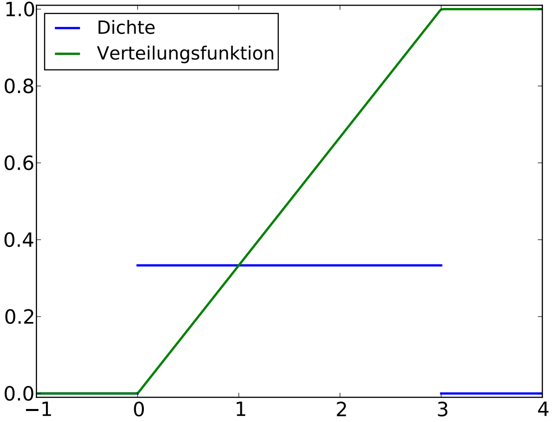
Dichte

$$f(x) = \begin{cases} 0, & x < c, \\ \frac{1}{d-c}, & c \leq x \leq d \\ 0, & x > d \end{cases}$$

Verteilungsfunktion

$$F(x) = \begin{cases} 0, & x < c \\ \frac{x-c}{d-c}, & c \leq x \leq d \\ 1, & x > d. \end{cases}$$

Dichte und Verteilungsfunktion der Gleichverteilung auf $[0, 3]$



Exponentialverteilung

- Eine Zufallsvariable X ist *exponentialverteilt* zum Parameter $\lambda > 0$, wenn für je zwei positive Zahlen $a < b$ gilt

$$P(a < X \leq b) = \lambda \int_a^b e^{-\lambda \cdot x} dx.$$

- Das Integral kann man ausrechnen

$$P(a < X \leq b) = e^{-\lambda \cdot a} - e^{-\lambda \cdot b}$$

- Erwartungswert und Varianz

$$E(X) = \frac{1}{\lambda}, \qquad Var(X) = \frac{1}{\lambda^2}$$

Dichte und Verteilungsfunktion der Gleichverteilung auf $[c, d]$

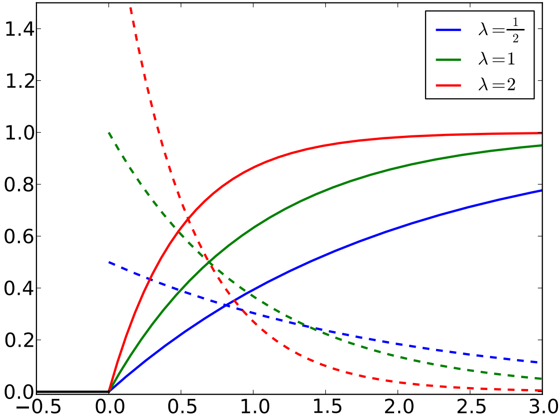
Dichte

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Verteilungsfunktion

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Dichte und Verteilungsfunktion von Exponentialverteilungen



Beispiel: Wartezeiten

Die Exponentialverteilung wird benutzt, um Wartezeiten zu modellieren. In diesem Fall ist der Parameter λ der Kehrwert der mittleren Wartezeit.

Beispiel: Glühbirne

- Eine Glühbirne hält im Schnitt 750h
- X = Lebensdauer der Glühbirne
- X exponentialverteilt zum Parameter $\lambda = 1/750 = 0.001333$
- Die Wahrscheinlichkeit, dass die Glühbirne mindestens 1000h hält, ist gleich

$$\begin{aligned} P(X \geq 1000) = 1 - P(X < 1000) &= 1 - F(1000) \\ &= 1 - (1 - e^{-.00133 \cdot 1000}) = e^{-1.33} = 0.26 \end{aligned}$$

- Die Wahrscheinlichkeit, dass eine Glühbirne mindestens 1000h hält, ist 26%

Abschnitt 2.10

Mehrere Zufallsvariable

Gemeinsame Verteilung für diskrete Zufallsvariable

$X_1, X_2, \ldots, X_n$ seien unabhängig und diskret. Dann

$$\begin{aligned} P(X_1 = k_1, X_2 = k_2, \ldots, X_n = k_n) \\ &= P(X_1 = k_1) \cdot P(X_2 = k_2) \cdots P(X_n = k_n) \end{aligned}$$

Gemeinsame Verteilung für stetige Zufallsvariable

$X_1, X_2, \ldots, X_n$ seien unabhängig. Die Verteilung von X_j habe die Dichte f_j. Dann

$$\begin{aligned}
&P(a_1 < X_1 \leq b_1,\, a_2 < X_2 \leq b_2, \ldots, a_n < X_n \leq b_n) \\
&\qquad = P(a_1 < X_1 \leq b_1) \cdot P(a_2 < X_2 \leq b_2) \cdots P(a_n < X_n \leq b_n) \\
&\qquad = \int_{a_1}^{b_1} f_1(x)\, dx \cdot \int_{a_2}^{b_2} f_2(x)\, dx \cdots \int_{a_n}^{b_n} f_n(x)\, dx \\
&\qquad = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \ldots \int_{a_n}^{b_n} f(x_1, x_2, \ldots, x_n)\, dx_n\, dx_2 \ldots dx_1
\end{aligned}$$

für

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) \cdot f_2(x_2) \cdots f_n(x_n)$$

Man bezeichnet f als Dichte der gemeinsamen Verteilung.

Das schwache Gesetz der großen Zahl

- "Mit ausreichend vielen Messwiederholungen lässt sich jede Genauigkeit erreichen."
- Präziser: $X_1, \ldots, X_n$ unabhängig, alle mit derselben Verteilung
- $\mu = E(X_1) = \cdots = E(X_n)$ und
  $\sigma^2 = Var(X_1) = \cdots = Var(X_n)$
- $Y = \frac{1}{n}(X_1 + \cdots + X_n)$
- Dann $E(Y) = \mu$ und die Streuung von Y beträgt

$$\sigma_Y = \frac{\sigma}{\sqrt{n}}$$