11. Grundbegriffe

11.1 Wahrscheinlichkeitsdichte und Verteilungsfunktion

11.2 Erwartungswert

11.3 Varianz und Momente von Verteilungen

11.4 Weitere Parameter von Verteilungen

11.5 Verallgemeinerung auf mehrere Variablen, Kovarianz

11.6 Beschreibung von Daten

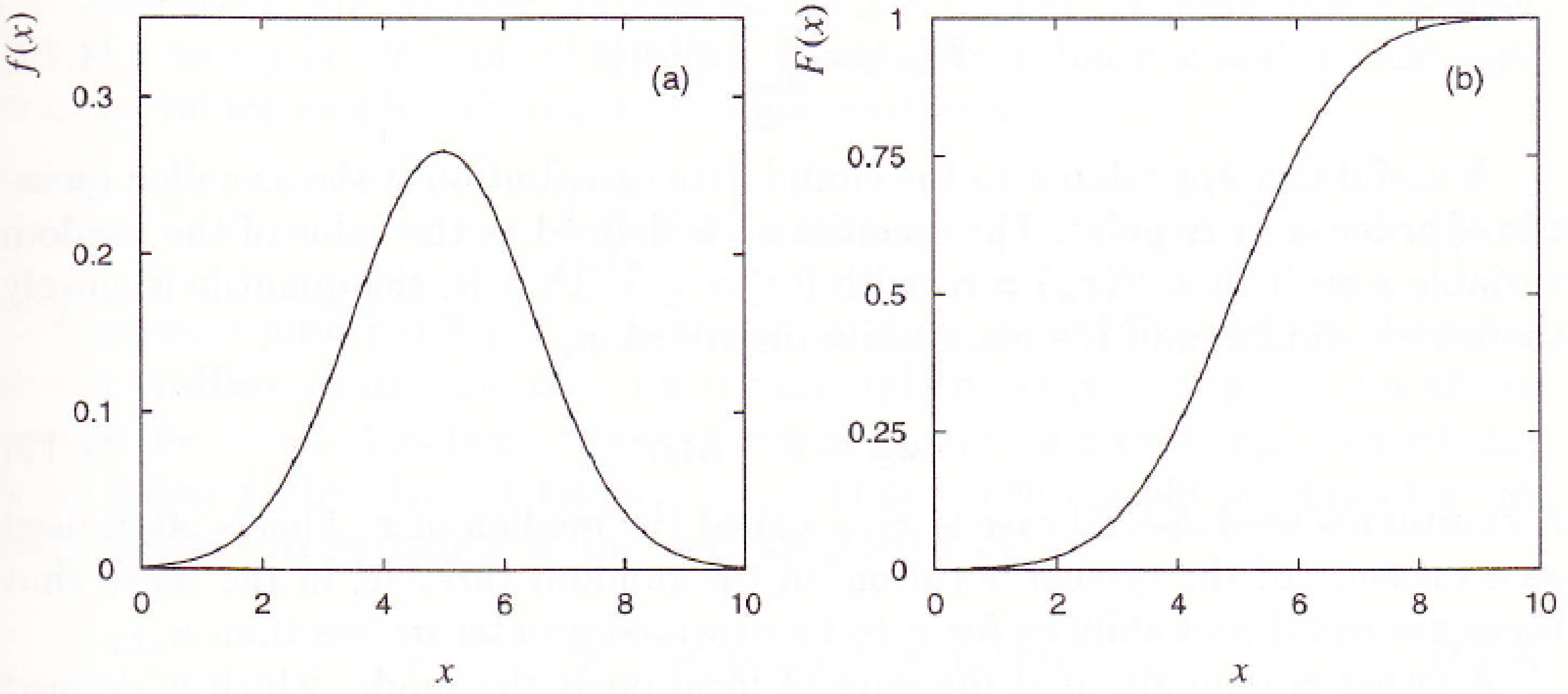

Fig. 1.3 (a) A probability density function $f(x)$. (b) The corresponding cumulative distribution function $F(x)$.

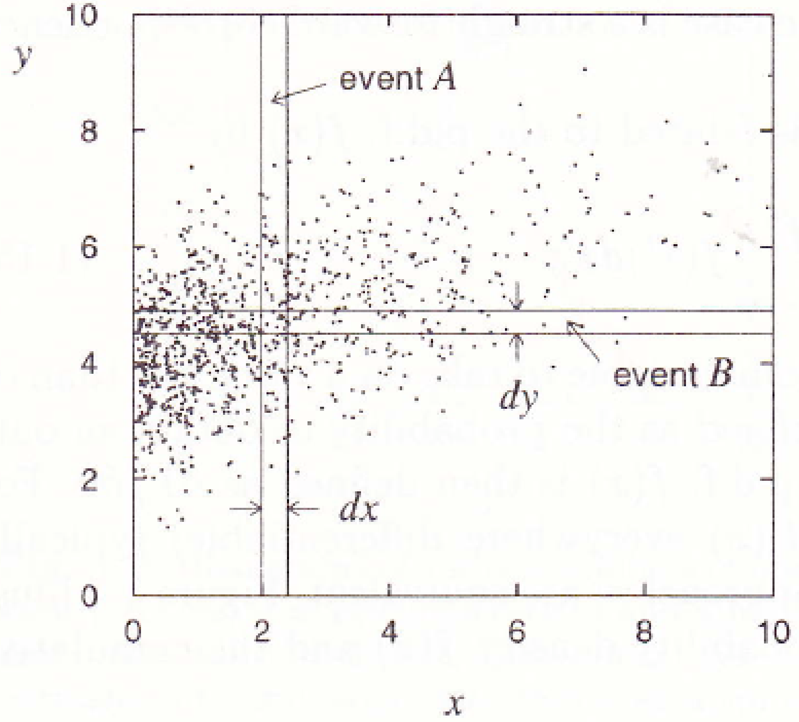

Fig. 1.4 A scatter plot of two random variables x and y based on 1000 observations. The probability for a point to be observed in the square given by the intersection of the two bands (the event $A \cap B$) is given by the joint p.d.f. times the area element, $f(x, y)dxdy$.

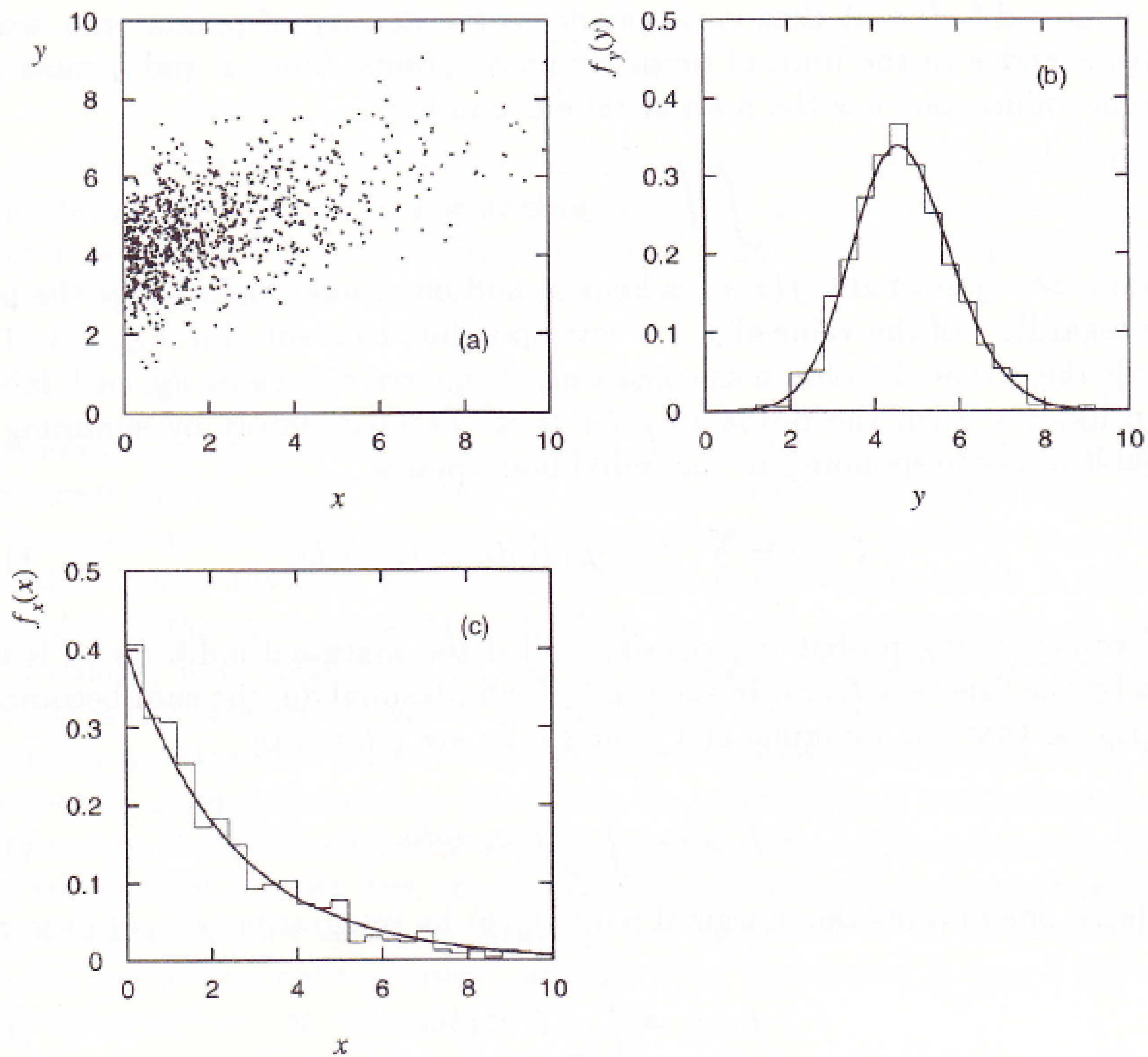

Fig. 1.5 (a) The density of points on the scatter plot is given by the joint p.d.f. $f(x, y)$. (b) Normalized histogram from projecting the points onto the y axis with the corresponding marginal p.d.f. $f_y(y)$. (c) Projection onto the x axis giving $f_x(x)$.

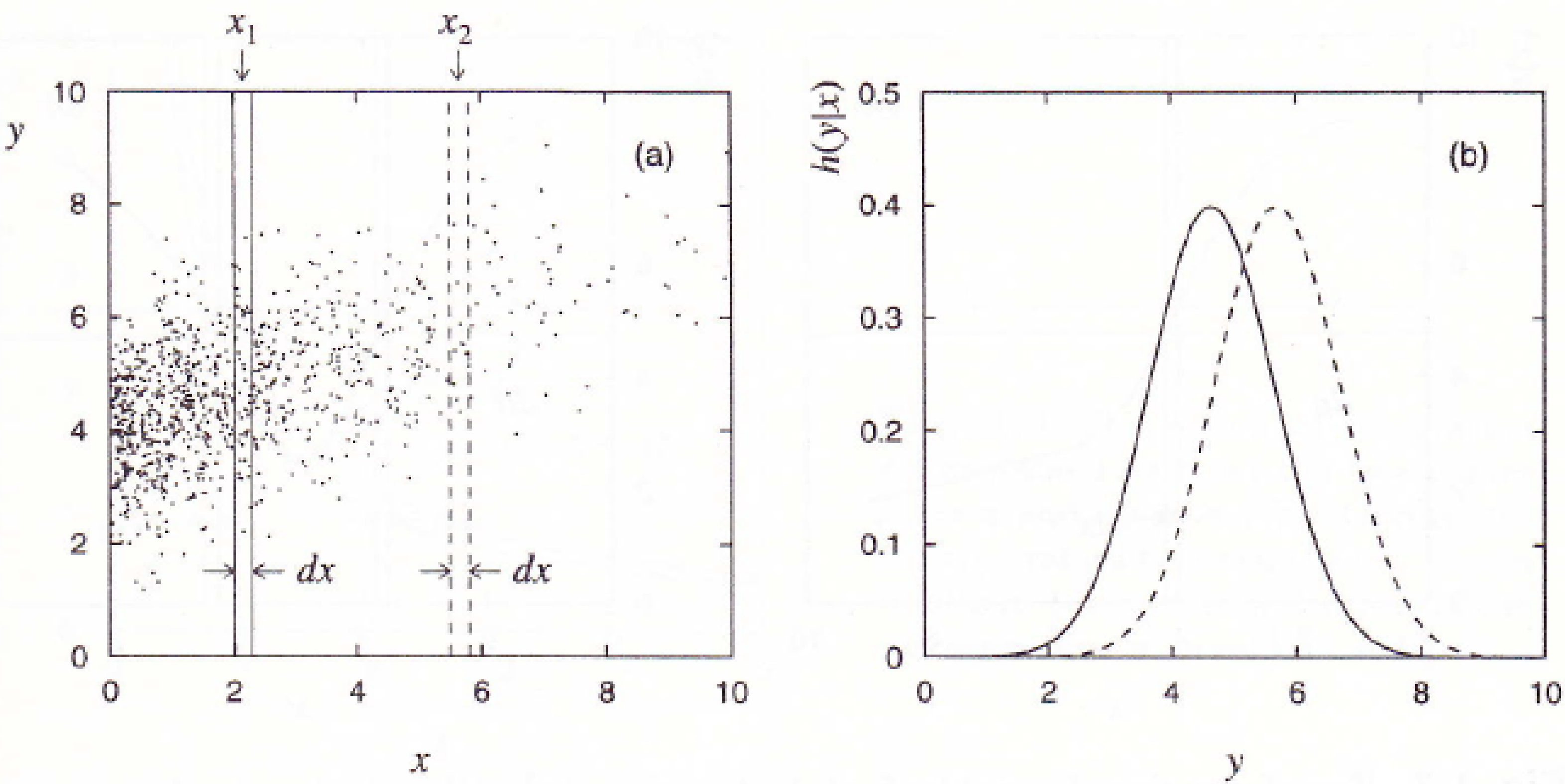

Fig. 1.6 (a) A scatter plot of random variables x and y indicating two infinitesimal bands in x of width dx at x_1 (solid band) and x_2 (dashed band). (b) The conditional p.d.f.s $h(y|x_1)$ and $h(y|x_2)$ corresponding to the projections of the bands onto the y axis.

Verallgemeinerung auf n Variablen

Variablen: $x_1, x_2, x_3, \ldots\ldots, x_n$

Verteilungsfkt: $F(x_1, x_2, x_3, \ldots X_n) = P(x_1 < X_1,\ x_2 < X_2, \ldots\ldots, x_n < X_n)$

Erwartungswerte:

$$E\{h(x_1, x_2, \ldots, x_n)\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{\infty} h(x_1, \ldots x_n)\, f(x_1, \ldots, x_n)\, dx_1\, d \ldots dx$$

$$E(x_r) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{\infty} x_r \cdot f(x_1, x_2, \ldots, x_n)\, dx_1 \cdot dx_2 \cdot \ldots \cdot dx_n$$

Variablen sind **unabhängig**, wenn:

$$f(x_1, x_2, \ldots, x_n) = g_1(x_1) \cdot g_2(x_2) \cdot \ldots\ldots \cdot g_n(x_n)$$

Zwischen allen Variablenpaaren (x_i, x_j) können **Kovarianzen** berechnet werden:

$$c_{ij} = \mathrm{cov}(x_i, x_j) = E\{(x_i - \mu_i)(x_j - \mu_j)\}$$

Kovarianzmatrix:

$$C := \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & & \\ \vdots & & \ddots & \\ c_{n1} & & & c_{nn} \end{pmatrix}$$

Diagonale: $c_{ii} = \sigma^2(x_i)$

Symmetrisch: $c_{ij} = c_{ji}$

Kurzschreibweise: $\vec{x} = (x_1, x_2, \ldots x_n)$ $\quad \vec{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)$

$$C = E\{(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^T\}$$

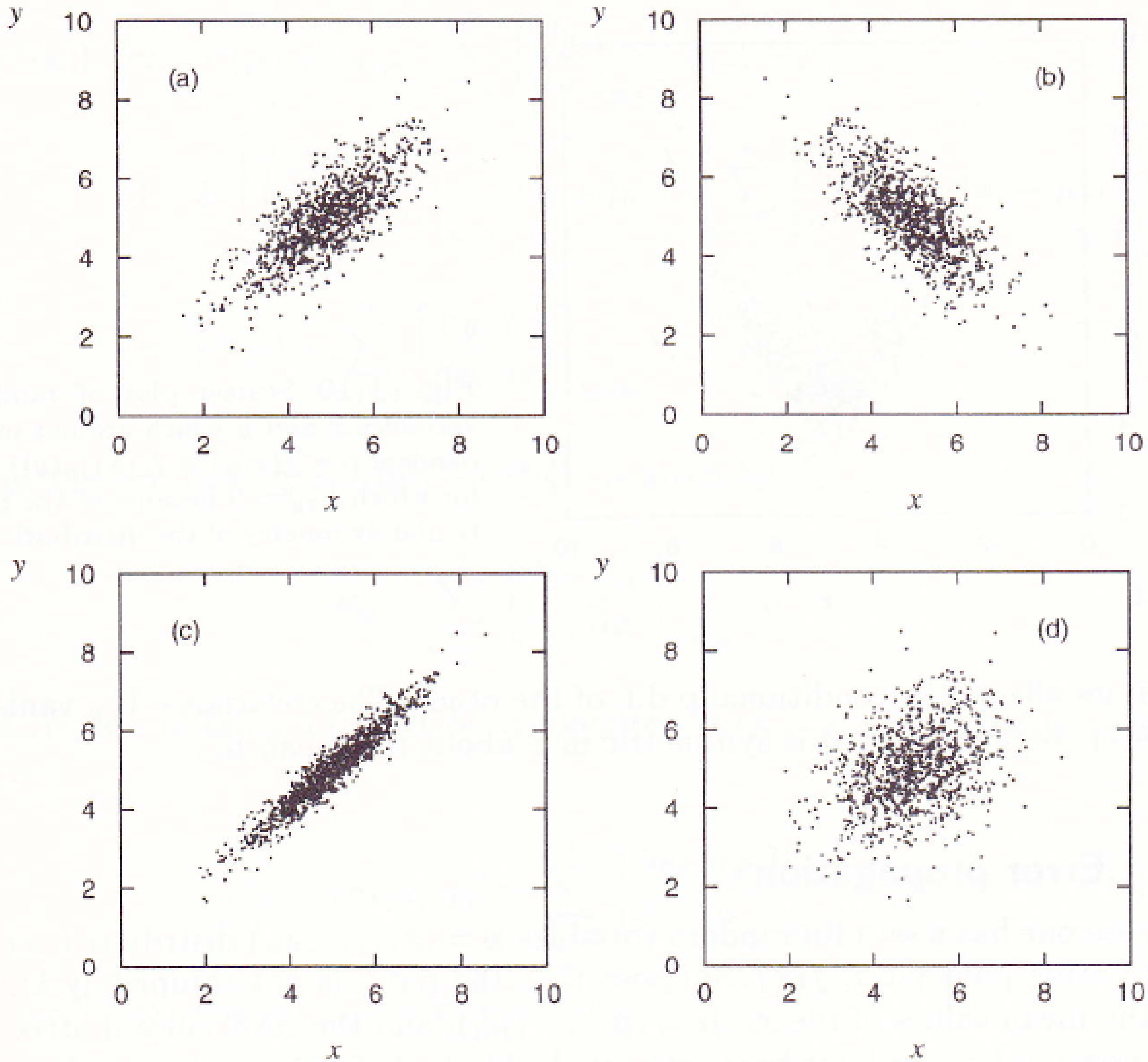

Fig. 1.9 Scatter plots of random variables x and y with (a) a positive correlation, $\rho = 0.75$, (b) a negative correlation, $\rho = -0.75$, (c) $\rho = 0.95$, and (d) $\rho = 0.25$. For all four cases the standard deviations of x and y are $\sigma_x = \sigma_y = 1$.