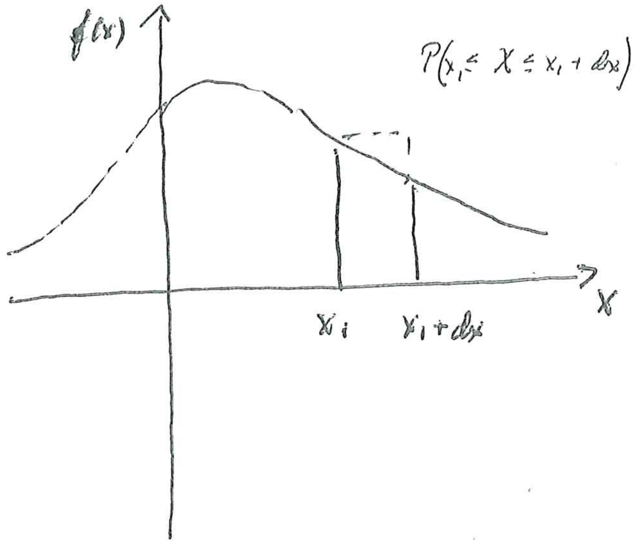
Multivariate distributions

Univariate (cont. distributions)

Cumulative distribution function: $F(x) = P(X \leq x)$



$P(x_1 \leq X \leq x_1 + dx) = F(x_1 + dx) - F(x_1) \approx f(x_1)\,dx$

Probability density function

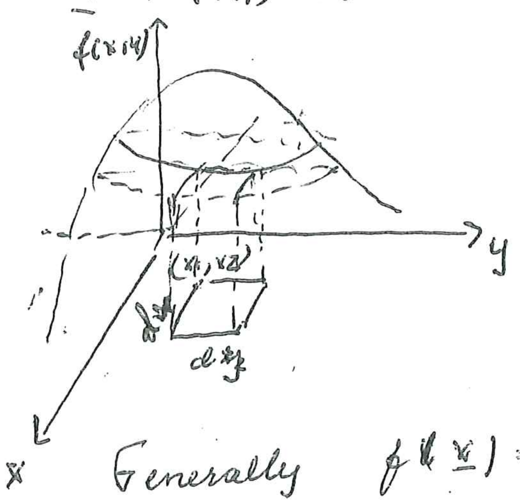
$f(x) = \frac{dF}{dx}$, $f(x) \geq 0, \forall x$, $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Multivariate

$\underline{X} = (X_1, X_2, \ldots, X_p)^T$

$F(\underline{x}) = P(\underline{X} \leq \underline{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_p \leq x_p)$

$\underline{X} = (X_1, X_2)^T$



$P(x_1 \leq X_1 \leq x_1 + dx, x_2 \leq X_2 \leq x_2 + dy)$

$\approx f(x_1, x_2)\,dx\,dy$

and $f(\underline{x}) = f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2}$.

Generally $f(\underline{x}) = \frac{\partial^p F(\underline{x})}{\partial x_1 \cdots \partial x_p}$ and $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\underline{x})\,dx_1, \ldots, dx_p = 1$

and $F(\underline{x}) = \int_{-\infty}^{x_p} \cdots \int_{-\infty}^{x_1} f(u_1, u_2, \ldots, u_p)\,du_1 \cdots du_p$.

Let
$$\underline{X} = [\underline{X}_1, \underline{X}_2]^T, \quad \underline{X}_1 \in R^k; \quad X_2 \in R^{p-k}$$

$$F_{X_1}(\underline{x}_1) = P(\underline{X}_1 \le \underline{x}_1) = F(x_{11}, \dots, x_{1k}, \infty \dots \infty)$$
$$= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(u_1 \cdots u_p)\, du_{k+1} \cdots du_p \right) du_1 \cdots du_k$$

is called the marginal cumulative density function for $\underline{X}_1$

The joint density function ~~is given by~~ for $x_{11}, \dots, x_{1k}$ is given by $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\underline{x}_1, \underline{x}_2)\, d\underline{x}_2$.

<u>Conditional distributions.</u>

Let X_1 and X_2 be two random variables
The conditional probability density function of X_2 given $X_1 = x_1$ is given as

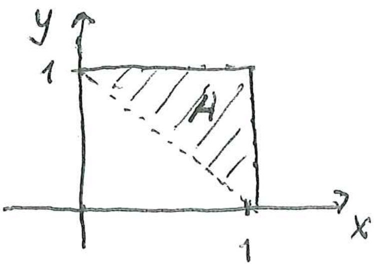
$$f(x_2 | x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)} \quad \text{provided that}$$

$$f_{X_1}(x_1) > 0$$

Examples $[X, Y]^T$ has pdf

$$f(x,y) = \begin{cases} 6xy^2, & 0 < x < 1, \ 0 < y < 1 \\ 0 & \text{else} \end{cases}$$

$P(X + Y \ge 1)$?



$$P([X, Y]^T \in A)$$

$$\int_0^1 \int_0^1 6xy^2\, dx\, dy = \int_0^1 6x \left[\frac{y^3}{3}\right]_0^1 dx = \int_0^1 2x\, dx = \left[\frac{x^2}{2}\right]_0^1 = 1$$

$$P(X+Y \geq 1) = \iint_A f(x,y)\,dx\,dy = \int_0^1 \left(\int_{1-x}^1 6xy^2\,dy\right) dx$$

$$= \int_0^1 6x \left[\frac{y^3}{3}\right]_{1-x}^1 dx = \int_0^1 2x\left(1-(1-x)^3\right) dx$$

$$= \int_0^1 2x\,dx - \int_0^1 2x(1-x)^3\,dx \overset{u=1-x}{=} 1 - \int_0^1 2(1-u)u^3\,du$$

$$= 1 - \left(\frac{1}{2} - \frac{2}{5}\right) = \frac{9}{10}$$

Marginal density

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 6xy^2\,dy = 2x\left[y^3\right]_0^1 = 2x, \quad 0<x<1$$

$$P\left(\frac{1}{2} < X < \frac{3}{4}\right) = \int_{\frac{1}{2}}^{3/4} 2x\,dx = \left[x^2\right]_{1/2}^{3/4} = \frac{9}{16} - \frac{1}{4} = \frac{5}{16}$$

Conditional density

$$f(y \mid X=x) = \frac{f(y,x)}{f_X(x)} = \frac{6xy^2}{2x} = 3y^2, \quad 0<y<1$$

<u>In general</u> let $\underline{X} = [\underline{X}_1, \underline{X}_2]^T$, $\underline{X}_1 \in R^k$, $\underline{X}_2 \in R^{p-k}$

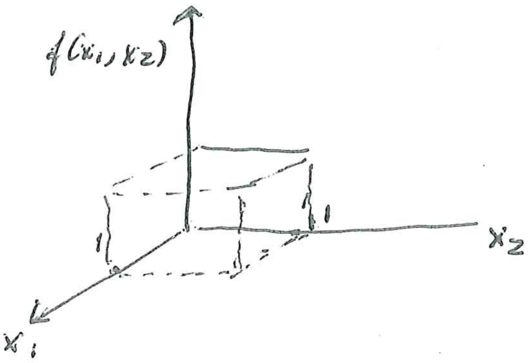
Then $$f(\underline{x}_2 \mid \underline{X}_1 = \underline{x}_1^0) = \frac{f(\overbrace{\underline{x}_1, \underline{x}_2}^{x_1, \dots, x_p})}{f(\underline{x}_1^0)}$$

$\underline{X}_1$ and $\underline{X}_2$ are independent if:

$$f(\underline{x}_1, \underline{x}_2) = f(\underline{x}_1)\cdot f(\underline{x}_2) = f(\underline{x}_1)\cdot f(\underline{x}_2 \mid \underline{X}_1 = \underline{x}_1), \quad \forall(\underline{x}_1, \underline{x}_2)$$

$$\Longleftrightarrow f(\underline{x}_2 \mid \underline{X}_1 = \underline{x}_1) = f(\underline{x}_2)$$

Joint and marginal distributions



$$f(x_1, x_2) = 1, \quad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1$$

$$f_{X_1}(x_1) = \int_0^1 1 \, dx_2 = 1, \quad 0 \le x_1 \le 1$$

$$\shortparallel$$

$$f_{X_2}(x_2)$$

$$\Rightarrow \quad f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

Let

$$f(x_1, x_2) = 1 + \alpha(2x_1 - 1)(2x_2 - 1), \quad 0 < x_1 < 1, \quad 0 < x_2 < 1, \quad |\alpha| < 1$$

$$\Rightarrow \quad f_{X_1}(x_1) = \int_0^1 (1 + \alpha(2x_1 - 1)(2x_2 - 1)) \, dx_2$$

$$= 1 + \alpha(2x_1 - 1)\left[x_2^2 - x_2\right]_0^1 = 1, \quad 0 < x_1 < 1$$

$$f_{X_2}(x_2) = 1, \quad 0 < x_2 < 1$$

$$f(x_1, x_2) \neq f_{X_1}(x_1) \cdot f_{X_2}(x_2) \quad \Rightarrow \text{dependence}.$$

Moments

X random variable (continuous)

$$E[X] = \int_{-\infty}^{\infty} x \, f(x) \, dx \qquad \left(\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow[n \to \infty]{a.s.} E[X] \right)$$

Let

$$\underline{X} = [X_1, X_2, \dots, X_p]$$

$$E[\underline{X}] = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_p] \end{bmatrix} = \begin{bmatrix} \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1) \, dx_1 \\ \vdots \\ \int_{-\infty}^{\infty} x_p f_{X_p}(x_p) \, dx_p \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix}$$

Some rules, providing the expectations exist

$$E[\alpha \underline{X} + \beta \underline{Y}] = \alpha E[\underline{X}] + \beta E[\underline{Y}] \qquad \text{linearity.}$$

$\underline{A}_{(q \times p)}$: $\quad E[\underline{A}\,\underline{X}] = \underline{A}\, E[\underline{X}]$

The i-th element in $E[\underline{A}\,\underline{X}]$ is $E\left[\sum_{k=1}^{p} a_{ik} X_k\right] = \sum_{k=1}^{p} a_{ik} E[X_k]$

where a_{ik} is the element on place ik in $\underline{A}$

In general let $\underline{X}$ be a matrix with random variable X_{ij} on place ij. Let $\underline{A}$ and $\underline{B}$ be matrices such that $\underline{A}\,\underline{X}\,\underline{B}$ is defined. Then $E[\underline{A}\,\underline{X}\,\underline{B}] = \underline{A}\, E[\underline{X}]\, \underline{B}$

Let $\underline{X}$ and $\underline{Y}$ be independent random vectors. Then

$$E[\underline{X}\,\underline{Y}^T] = E\left[\begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} (Y_1, \dots, Y_q)\right] = (E\underline{X})(E\underline{Y})^T$$

~~alle komponenter i $\underline{X}$ er uavhengige av komponenter $\underline{Y}$~~

The Covariance matrix $\underline{\Sigma} = Var[\underline{X}]$

__Univariate.__ $\quad Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$

$Y = X$: $\quad Cov[X,X] = E[(X - E[X])^2] = Var[X]$

Let $\underline{X}$ be a random vector, $\underline{X} = [X_1, \dots, X_p]^T$

$$Cov[\underline{X}] = E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^T] = Var[\underline{X}] = \underline{\Sigma}$$

$$= E[\underline{X}\,\underline{X}^T] - E[\underline{X}\, E[\underline{X}]^T] - E[E[\underline{X}]\,\underline{X}^T] + E[E[\underline{X}] \cdot E[\underline{X}]^T]$$

$$= E[\underline{X}\,\underline{X}^T] - E[\underline{X}] \cdot E[\underline{X}]^T$$