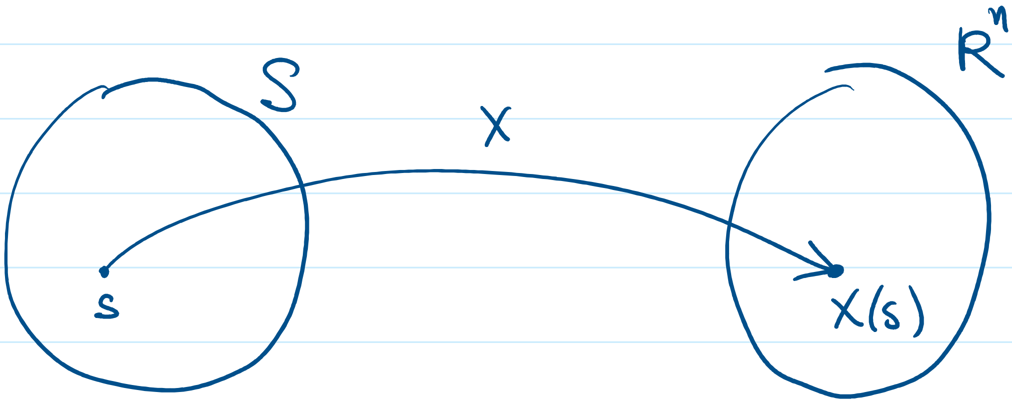
4.6 Multivariate distributions

A Random vector denoted by $X = (X_1, \dots, X_n)$ is a function that maps elements of a sample space S into $\mathbb{R}^n$.



$$X(s) = (X_1(s), \dots, X_n(s)) \in \mathbb{R}^n$$

Now, we may specify probabilities, for $A \subset \mathbb{R}^n$

$$P(X \in A) := P(s \in S : X(s) \in A).$$

if X takes countably many values, then we say X is a discrete random vector, which is associated with a joint probability mass function (pmf) such that

$$\boxed{f(x) = f(x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n)}$$

Hence for $A \subset \mathbb{R}^n$

$$P(X \in A) = \sum_{x \in A} P(X = x)$$

$$= \sum_{x \in A} f(x).$$

We say $X = (X_1, \dots, X_n)$ is a continuous Random vector, if it is associated with a joint probability density function

$$f(x) = f(x_1, \dots, x_n)$$

such that for $A \subset \mathbb{R}^n$

$$\boxed{\begin{aligned} P(X = (X_1 \dots X_n) \in A) &= \int_A f(x)\, dx \\ &= \iint\dots\int_{(x_1 \dots x_n) \in A} f(x_1 \dots x_n)\, dx_1 \dots dx_n \end{aligned}}$$

Expectation

Let f be joint pdf/pmf for R.V. $X \in \mathbb{R}^n$.
Let $g : \mathbb{R}^n \to \mathbb{R}$ be a function on $\mathbb{R}^n$.
Then $g(X)$ is a (real-valued) Random variable and

$$\boxed{\mathbb{E}\, g(X) = \begin{cases} \displaystyle\int_{\mathbb{R}^n} g(x) f(x)\, dx & \text{if } X \text{ cont} \\ \displaystyle\sum_{x \in \mathbb{R}^n} g(x) f(x)\, dx & \text{if } X \text{ discrete} \end{cases}}$$

From joint to marginal distributions.

Suppose $(X_1, \dots, X_n)$ be a random vector in $\mathbb{R}^n$
and

$(X_1 \dots X_n) \sim f_X$ (joint pdf or pmf)

Let $1 \le k < n$.

Then the (marginal) distribution of $(X_1 \dots X_k)$
has a the pdf / pmf :

$$\boxed{f_{X_1 \dots X_k}(x_1, \dots, x_k) = \begin{cases} \displaystyle\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f_{X_1 \dots X_n}(x_1, \dots, x_n)\, dx_{k+1} \dots dx_n & \text{for continuous case} \\ \text{or} & \\ \displaystyle\sum_{x_{k+1}} \cdots \sum_{x_n} f_{X_1 \dots X_n}(x_1, \dots, x_n). & \text{for discrete case.} \end{cases}}$$

(a.d) Conditional distribution

of $(X_{k+1} \dots X_n)$ given $(X_1 \dots X_k) = (x_1, \dots, x_k)$, provided that $f_{(X_1 \dots X_k)}(x_1 \dots x_k) > 0$ admits the conditional pmf / pdf of the form

$$\boxed{f_{X_{k+1} \dots X_n \mid X_1 \dots X_k}(x_{k+1}, \dots, x_n \mid x_1 \dots x_k) = \frac{f(x_1 \dots x_k, x_{k+1} \dots x_n)}{f(x_1, \dots, x_k)}.}$$

Remark

① The joint pdf / pmf admits the factorization:

$$f_{X_1 \dots X_n}(x_1 \dots x_n) = f_{X_1 \dots X_k}(x_1 \dots x_k) \, f_{(X_{k+1} \dots X_n \mid X_1 \dots X_k)}(x_{k+1} \dots x_n \mid x_1 \dots x_k)$$

② For mixed continuous / discrete random vector the above factorization form continues to hold with suitable denotation.

For example if $\begin{cases} X, Y & \text{discrete} \\ Z, T & \text{continuous} \end{cases}$

Let

$$\boxed{f_{XYZT}(x, y, z, t) = f_{XY}(x, y)\, f_{ZT|XY}(z, t \mid x, y)}$$

pmf

conditional pdf

$$= f_{ZT}(z, t)\, f_{XY|ZT}(x, y \mid z, t)$$

Then, for any $A, B, C, D \subset \mathbb{R}$,

$$\boxed{\begin{aligned} &P(X \in A, Y \in B, Z \in C, T \in D) \\ &= \sum_{x \in A} \sum_{y \in B} \int_D \int_C f_{XYZT}(x, y, z, t)\, dz\, dt \end{aligned}}$$

Summation over discrete values
integration over continuous values

Example Multinomial distribution

Experiments: m indep. trials

n possible outcomes with

"cell probabilities" $(p_1, \dots, p_n)$, $\sum_{i=1}^{n} p_i = 1$.

$(X_1, \dots, X_n)$ represent the counts of each of the n outcomes

$X_1 + \dots + X_n = m$

joint pmf: for $x_1, \dots, x_n \in \mathbb{N}^n$, $\sum_{j=1}^{n} x_j = m$

$$\boxed{\begin{aligned} f(x_1, \dots, x_n) &= \binom{m}{x_1 \dots x_n} p_1^{x_1} \dots p_n^{x_n} \\ &= \frac{m!}{x_1! \dots x_n!} p_1^{x_1} \dots p_n^{x_n} \end{aligned}}$$

otherwise, $f(x_1 \dots x_n) = 0$.

Remarks

- By Multinomial formula

$$\sum_{\substack{x_1 \dots x_n \\ x_1+\dots+x_n=m}} \binom{m}{x_1 \dots x_n} p_1^{x_1} \cdots p_n^{x_n}$$

$$= (p_1+\dots+p_n)^m$$

$$= 1$$

- Marginal distributions

$$f_{X_n}(x_n) = \sum_{x_1+\dots+x_{n-1}=m-x_n} \frac{m!}{x_1!\dots x_{n-1}!\,x_n!} p_1^{x_1} \cdots p_{n-1}^{x_{n-1}} p_n^{x_n}$$

$$= p_n^{x_n} \frac{m!}{x_n!\,(m-x_n)!} \sum_{\substack{x_1+\dots+x_{n-1} \\ =m-x_n}} \frac{(m-x_n)!}{x_1!\dots x_{n-1}!} p_1^{x_1} \cdots p_{n-1}^{x_{n-1}}$$

$$= \frac{m!}{x_n!\,(m-x_n)!} p_n^{x_n} (p_1+\dots+p_{n-1})^{m-x_n}$$

$$= \frac{m!}{x_n!\,(m-x_n)!} p_n^{x_n} (1-p_n)^{m-x_n}$$

Hence $X_n \sim \text{binomial}(m, p_n)$.

- Conditional distributions

$$f_{X_1 \dots X_{n-1} \mid X_n}(x_1 \dots x_{n-1} \mid x_n) = \frac{f_{X_1 \dots X_n}(x_1 \dots x_n)}{f_{X_n}(x_n)}$$

$$= \frac{\frac{m!}{x_1! \dots x_n!} p_1^{x_1} \cdots p_{n-1}^{x_{n-1}} p_n^{x_n}}{\frac{m!}{x_n!(m-x_n)!} (1-p_n)^{m-x_n} p_n^{x_n}}$$

$$= \frac{(m-x_n)!}{x_1! \dots x_{n-1}!} \left(\frac{p_1}{1-p_n}\right)^{x_1} \cdots \left(\frac{p_{n-1}}{1-p_n}\right)^{x_{n-1}}$$

Hence,

$$X_1 \dots X_{n-1} \mid X_n = x_n \sim \text{Multinomial}\left(m - x_n \text{ trial, and cell probabilities } \left(\frac{p_1}{1-p_n}, \dots, \frac{p_{n-1}}{1-p_n}\right)\right).$$

- We saw a kind of "invariance" property by marginalization and conditioning

This holds for quite a few families of distributions (often in the exponential families)
For continuous distributions, one such family is the multivariate normal.

Mutual independence

Recall:

Def. Let (X,Y) be a bivariate Random vector, with joint pdf / pmf $f_{XY}(x,y)$ and marginal pdf / pmf $f_X(x)$ and $f_Y(y)$.

Then X and Y are called independent variables, if for every $x,y \in \mathbb{R}$

$$\boxed{f_{XY}(x,y) = f_X(x)\, f_Y(y)}.$$

For multivariate Random vector

Def. • Let $(X_1,\dots,X_n)$ be a Random vector with joint pdf/pmf $f(x_1,\dots,x_n)$

• Let $f_{X_i}(x_i)$ be the marginal pdf of X_i

Then $X_1, X_2, \dots, X_n$ are mutually independent Random variables if, for every $(x_1,\dots,x_n)$

$$\boxed{f(x_1,\dots,x_n) = \prod_{i=1}^{n} f_{X_i}(x_i).}$$

Remark.

- We allow in the above definition the domain of each of X_i to be (subset) of $\mathbb{R}^{n_i}$ for some $n_i \in \mathbb{N}$.
(in other words, each X_i may be itself a random vector)

- Many properties established for bivariate random vector can be extended naturally to the multivariate setting.

Theorem if $X_1, \dots, X_n$ are (mutually independent) R.V.'s
Let $g_1(x_1), \dots, g_n(x_n)$ be real-valued functions on the domain of $X_1, \dots, X_n$, respectively.
Then

$$\boxed{\mathbb{E}\left(g_1(X_1) \cdots g_n(X_n)\right) = \mathbb{E}(g_1(X_1)) \cdots \mathbb{E}(g_n(X_n))}.$$

Theorem if $X_1, \dots, X_n$ are (mutual independent) R.V.'s
Let $M_{X_i}(t)$ be the mgf for X_i $(i=1,\dots,n)$
Then $Z := X_1 + \dots + X_n$ has the mgf

$$\boxed{M_Z(t) = M_{X_1}(t) \cdots M_{X_n}(t).}$$

in particular, if $X_1, \dots, X_n$ are i.i.d., then

$$M_Z(t) = \left(M_{X_1}(t)\right)^n.$$

From these theorems, we can obtain the following easy facts:

Consequences:

① if $X_i \sim \text{gamma}(\alpha_i, \beta)$, X_i are indep.
Then

$$X_1 + \dots + X_n \sim \text{gamma}(\alpha_1 + \dots + \alpha_n, \beta).$$

② if $X_i \overset{\text{indep.}}{\sim} N(\mu_i, \sigma_i^2)$
Let $a_i, b_i \in \mathbb{R}$ Then

$$Z := \sum_{i=1}^n (a_i X_i + b_i) \sim N\left(\sum_{i=1}^n (a_i \mu_i + b_i), \sum_{i=1}^n a_i^2 \sigma_i^2\right)$$

Recognizing independence

Lemma (extending from the bivariate case in Sec 4.2b)

Suppose $(X_1, \dots, X_n) \sim f_{X_1 \dots X_n}$

$X_1 \dots X_n$ are mutually independent random variables if and only if its joint pdf / pmf "factorizes":

$$f_{X_1 \dots X_n}(x_1, \dots, x_n) = g_1(x_1) \dots g_n(x_n)$$

for some functions $g_1, \dots, g_n$.

Proof. Similar to the bivariate case.

Theorem if $X_1, \dots, X_n$ are mutually indep.
Let $g_i(x_i)$ be a function of x_i, $i=1,\dots,n$
Then

$$g_1(X_1), \dots, g_n(X_n) \text{ are also mutually indep.}$$

Proof. Follows directly from the factorization of the joint pdf / pmf.

The change of variable formula can be extended to the multivariate case too, see text book.