1. DISTRIBUTION THEORY

2. Clearly independence \( \iff \)
\[
P_{X_i|X_1 \ldots X_{i-1}, X_{i+1} \ldots X_r}(x_i|x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_r) = P_{X_i}(x_i). \quad i = 1, \ldots, r
\]
Moreover, we have:
\[
P_{X_i|X_2}(x_1|x_2) = P_{X_i}(x_1) \quad \text{for any partitioning of } X = \left( \begin{array} {c} X_1 \\ X_2 \end{array} \right) \text{ if } X_1, \ldots, X_r \text{ are independent.}
\]

1.7.4 Continuous multivariate distributions

Definition. 1.7.5

The random vector \((X_1, \ldots, X_r)^T\) is said to have a continuous multivariate distribution with PDF \(f(x)\) if
\[
P(X \in A) = \int \ldots \int_A f(x_1, \ldots, x_r) \, dx_1 \ldots dx_r
\]
for any measurable set \(A\).

Examples

1. Suppose \((X_1, X_2)\) have the trinomial distribution with parameters \(n, \pi_1, \pi_2\). Then the marginal distribution of \(X_1\) can be seen to be \(B(n, \pi_1)\) and the conditional distribution of \(X_2|X_1 = x_1\) is \(B(n - x_1, \frac{\pi_2}{1 - \pi_1})\).

   Outcome 1 \(\pi_1\)
   Outcome 2 \(\pi_2\)
   Outcome 3 \(1 - \pi_1 - \pi_2\)

Examples of Definition 1.7.5

1. If \(X_1, X_2\) have PDF
\[
f(x_1, x_2) = \begin{cases} 1 & 0 < x_1 < 1 \\ 0 & x_2 < 1 \\ 0 & \text{otherwise} \end{cases}
\]
The distribution is called uniform on \((0,1) \times (0,1)\).

   \(\text{Area} = 1\)

   Why?
1. DISTRIBUTION THEORY

2. Uniform distribution on unit disk:

\[
f(x_1, x_2) = \begin{cases} \frac{1}{\pi} & x_1^2 + x_2^2 < 1 \\ 0 & \text{otherwise.} \end{cases}
\]

3. Dirichlet distribution is defined by PDF

\[
f(x_1, x_2, \ldots, x_r) = \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_r + \alpha_{r+1})}{\Gamma(\alpha_1)\Gamma(\alpha_2)\ldots\Gamma(\alpha_r)\Gamma(\alpha_{r+1})} x_1^{\alpha_1-1}x_2^{\alpha_2-1}\cdots x_r^{\alpha_r-1}(1 - x_1 - \cdots - x_r)^{\alpha_{r+1}}
\]

for \( x_1, x_2, \ldots, x_r > 0 \), \( \sum x_i < 1 \), and parameters \( \alpha_1, \alpha_2, \ldots, \alpha_{r+1} > 0 \).

Recall joint PDF:

\[
P(X \in A) = \int_A \ldots \int_A f(x_1, \ldots, x_r) \, dx_1 \ldots dx_r.
\]

Note: the joint PDF must satisfy:

1. \( f(x) \geq 0 \) for all \( x \);
2. \( \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_r) \, dx_1 \ldots dx_r = 1. \)
Definition. 1.7.6

If $\mathbf{X} = \left( X_1, X_2 \right)$ has joint PDF $f(x) = f(x_1, x_2)$, then the marginal PDF of $X_1$ is given by:

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_r, x_{r+1}, \ldots, x_r) \, dx_{r+1} \cdots dx_{r+2} \cdots dx_r,$$

and for $f_{X_1}(x_1) > 0$, the conditional PDF is given by

$$f_{X_2|X_1}(x_2|x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)}.$$

Remarks:

1. $f_{X_2|X_1}(x_2|x_1)$ cannot be interpreted as the conditional PDF of $X_2|\{X_1 = x_1\}$ because $P(X_1 = x_1) = 0$ for any continuous distribution.

   Proper interpretation is the limit as $\delta \to 0$ in $X_2|X_1 \in B(x_1, \delta)$.

2. $f_{X_1}(x_1), f_{X_1|X_2}(x_1|x_2)$ are defined analogously.

Definition. 1.7.7 (Independence)

Continuous RV’s $X_1, X_2, \ldots, X_r$ are said to be (mutually) independent if their joint PDF satisfies:

$$f(x_1, x_2, \ldots, x_r) = f_1(x_1)f_2(x_2) \cdots f_r(x_r),$$

for some functions $f_1, f_2, \ldots, f_r$ and all $x_1, \ldots, x_r$

Remarks

1. Easy to check that if $X_1, \ldots, X_r$ are independent then each $f_i(x_i) = c_i f_{X_i}(x_i)$.

   Moreover $c_1, c_2, \ldots, c_r = 1$.

2. If $X_1, X_2, \ldots, X_r$ are independent, then it can be checked that:

   $$f_{X_1|X_2}(x_1|x_2) = f_{X_1}(x_1)$$

   for any partitioning of $\mathbf{X} = \left( X_1, X_2 \right)$.

Examples
1. If \((X_1, X_2)\) has the uniform distribution on the unit disk, find:

(a) the marginal PDF \(f_{X_1}(x_1)\)
(b) the conditional PDF \(f_{X_2|X_1}(x_2|x_1)\).

**Solution.** Recall

\[
f(x_1, x_2) = \begin{cases} 
\frac{1}{\pi} & x_1^2 + x_2^2 < 1 \\
0 & \text{otherwise}
\end{cases}
\]

(a) \(f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) \, dx_2\)

\[
= \int_{-\infty}^{-\sqrt{1-x_1^2}} 0 \, dx_2 + \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \frac{1}{\pi} \, dx_2 + \int_{\sqrt{1-x_1^2}}^{\infty} 0 \, dx_2
\]

\[
= 0 + \frac{x_2}{\pi} - \frac{\sqrt{1-x_1^2}}{\pi} + 0
\]

\[
\begin{cases} 
\frac{2\sqrt{1-x_1^2}}{\pi} & -1 < x_1 < 1 \\
0 & \text{otherwise}.
\end{cases}
\]

\(\Rightarrow\) **Drop** \(X_2\)

\(\Rightarrow\) \(X_1^2 < 1\)

\(\Rightarrow\) **i.e., a semi-circular distribution (ours has some distortion).**

(b) The conditional density for \(X_2|X_1\) is:

\[
f_{X_2|X_1}(x_2|x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)} = \left(\frac{x_2}{\sqrt{1-x_1^2}}\right) / \left(\frac{2\sqrt{1-x_1^2}}{\pi}\right)
\]

\[
\begin{cases} 
\frac{1}{2\sqrt{1-x_1^2}} & \text{for } -\sqrt{1-x_1^2} < x_2 < \sqrt{1-x_1^2} \\
0 & \text{otherwise},
\end{cases}
\]

which is uniform \(U(-\sqrt{1-x_1^2}, \sqrt{1-x_1^2})\).

(\text{Dotted line Fig 13})

2. If \(X_1, \ldots, X_r\) are any independent continuous RV’s then their joint PDF is:

\[
f(x_1, \ldots, x_r) = f_{X_1}(x_1)f_{X_2}(x_2) \cdots f_{X_r}(x_r).
\]
1. DISTRIBUTION THEORY

Figure 13: A graphic showing the conditional distribution and a semi-circular distribution.

E.g. If $X_1, X_2, \ldots, X_r$ are independent $\text{Exp}(\lambda)$ RVs, then

$$f(x_1, \ldots, x_r) = \prod_{i=1}^{r} \lambda e^{-\lambda x_i}$$

$$= \lambda^r e^{-\lambda \sum_{i=1}^{r} x_i},$$

for $x_i > 0 \quad i = 1, \ldots, n.$

3. Suppose $(X_1, X_2)$ are uniformly distributed on $(0, 1) \times (0, 1)$. That is,

$$f(x_1, x_2) = \begin{cases} 1 & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Claim: $X_1, X_2$ are independent.

Proof.

Let

$$f_1(x_1) = \begin{cases} 1 & 0 < x_1 < 1 \\ 0 & \text{otherwise}, \end{cases}$$

30
1. DISTRIBUTION THEORY

By defn

\[
f_2(x_2) = \begin{cases} 
1 & 0 < x_2 < 1 \\
0 & \text{otherwise,}
\end{cases}
\]

then \( f(x_1, x_2) = f_1(x_1)f_2(x_2) \), and the two variables are independent. \( \Box \)

4. (\( X_1, X_2 \)) is uniform on unit disk. Are \( X_1, X_2 \) independent? No.

**Proof.**
We know:

\[
f_{X_2|X_1}(x_2|x_1) = \begin{cases} 
\frac{1}{2\sqrt{1-x_1^2}} & -\sqrt{1-x_1^2} < x_2 < \sqrt{1-x_1^2} \\
0 & \text{otherwise},
\end{cases}
\]

i.e., \( U(-\sqrt{1-x_1^2}, \sqrt{1-x_1^2}) \).

On the other hand,

\[
f_{X_2}(x_2) = \begin{cases} 
\frac{2}{\pi} \sqrt{1-x_2^2} & -1 < x_2 < 1 \\
0 & \text{otherwise},
\end{cases}
\]

i.e., a semicircular distribution.

Hence \( f_{X_2}(x_2) \neq f_{X_2|X_1}(x_2|x_1) \), so the variables cannot be independent. \( \Box \)

**Definition.** 1.7.8

The joint CDF of RVS's \( X_1, X_2, \ldots, X_r \) is defined by:

\[
F(x_1, x_2, \ldots, x_r) = P\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\} \cap \ldots \cap \{X_r \leq x_r\}.
\]

**Remarks:**

1. Definition applies to RV's of any type, i.e., discrete, continuous, “hybrid”.

2. Marginal CDF of \( X_s = (X_1, \ldots, X_s)^T \) is \( F_{X_1}(x_1, \ldots, x_s) = F_X(x_1, \ldots, x_s, \infty, \infty, \ldots, \infty) \), for \( s < r \).

3. RV's \( X_1, \ldots, X_r \) are defined to be independent if

\[
F(x_1, \ldots, x_r) = F_{X_1}(x_1)F_{X_2}(x_2) \ldots F_{X_r}(x_r).
\]

4. The definitions above are completely general. However, in practice it is usually easier to work with the PDF/probability function for any given example.