Multivariate Distributions

1 Motivation

- Often when we are studying a particular topic we will take various measures that pertain to that topic
- From hre the goal is usually to examine these measures and see what they tell us about the topic we are studying
- Up until now we have only considered how we might approach one measure at a time, in the form of Random Variables
- But, it may be useful to understand how different measures relate to each other
- We can encapsulate this relationship by examining what we call the joint distribution between multiple Random Variables

2 Definitions & Properties

2.1 Discrete Joints PDFs

- While it is possible to have a joint distribution bewtween a mixture of discrete and continuous random variables, for this course we will focus on the joint distribution between random variables that are either all discrete random variables or all continuous random variables.

  Definition 1. *Let $X_1, X_2, ..., X_n$ all be discrete random variables, each with the support $S_1, ...S_n$, respectively. The the Joint probability distribution function (joint PDF, for short) is the function p such that*

$$p(x_1, ..., x_n) = \begin{cases} P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) & \textit{For } x_1 \in S_1, ..., x_n \in S_n \\ 0 & \textit{else} \end{cases}$$

- A few Notes about the joint PDF of Discrete Random Variables:
  - The statement $P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n)$ is equivalent to the statement $P((X_1 = x_1) \cap (X_2 = x_2) \cap ... \cap (X_n = x_n))$, where $(X_1 = x_1), ..., (X_n = x_n)$ are all seperate events
  - The PDF equals $P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n)$ only when $x_1 \in S_1, ..., x_n \in S_n$. This means that if a single x_i is not in the set S_i for some $i = 1, 2, 3, ..., n$, then the whole PDF evaluates to be 0.
- The Discerte Joint PDF has the following properties:
  Let $X_1, ..., X_n$ be discrete random variables, let $S_1, ..., S_n$ be their corresponding supports, and let p be their joint PDF. Then

1. $p(x_1, ..., x_n) \geq 0 \forall x_1 \in S_1, ..., \forall x_n \in S_n$
2. $\sum_{x_1 \in S_1} \sum_{x_2 \in S_2} ... \sum_{x_n \in S_n} p(x_1, ..., x_n) = 1$

2.2 Joint CDFs

- Like with singular Random variables, there are PDFs and CDFs
- the CDF will help us to easily define the join PDF of continuous random variables, just as it did in the singualr case

  Definition 2. *Let $X_1, X_2, ..., X_n$ all be random variables, each with the support $S_1, ...S_n$, respectively. The the Joint cumulative distribution function (joint CDF, for short) is the function F such that*

$$F(x_1, ..., x_n) = P(X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n)$$

- The Joint CDF has the following properties:
  Let $X_1, ..., X_n$ be random variables, let $S_1, ..., S_n$ be their corresponding supports, and let F be their joint CDF. Additionally, let $x_1, ..., x_n$ be real constant vaues. Then

  1. If any of the values input into the CDF is $-\infty$, the the result is 0. i.e.

$$\begin{aligned} F(-\infty, ..., -\infty) &= F(x_1, -\infty, ..., -\infty) \\ &= F(x_1, x_2, -\infty, ..., -\infty) \\ &\vdots \\ &= F(x_1, ..., x_{n-1}, -\infty) \\ &= 0 \end{aligned}$$

  2. $F(\infty, ..., \infty) = 1$

2.3 Joint Continuous PDFs

- Like in the univariate case (the case where we are only working with one random variable), we will define the joint continuous PDF in terms of the joint CDF

  Definition 3. *Let $X_1, X_2, ..., X_n$ all be continuous random variables, each with the support $S_1, ...S_n$, respectively. Let F be the joint CDF of $X_1, ..., X_n$. Then the Joint probability density function (joint CDF, for short) is the non-negative function f such that*

$$F(x_1, ..., x_n) = \int_{-\infty}^{x_n} ... \int_{-\infty}^{x_1} f(t_1, .., t_n) \mathrm{d}t_1 ... \mathrm{d}t_n$$

  for all real values $x_1, ..., x_n$ (assuming such a function exists).

- The Continuous Joint PDF has the following properties:
Let $X_1, ..., X_n$ be continuous random variables, let $S_1, ..., S_n$ be their corresponding supports, and let f be their joint PDF. Then
    1. $f(x_1, ..., x_n) \geq 0 \forall x_1, ..., x_n$
    2. $\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, .., x_n) \mathrm{d}x_1 ... \mathrm{d}x_n = 1$

- It is worht noting that sometimes we are considering problems of the form

$$P((X_1, X_2, ..., X_n) \in S)$$

Where S is a region of R^n.

- If $X_1, X_2, ..., X_n$ are discrete random variables, then the problem is fairly straightforward

- If $X_1, X_2, ...X_n$ are continuous, then we refer to the following theorem to solve the problem:

Theorem 1. *Let $X_1, X_2, ..., X_n$ be continuous random variables with joint pdf f and let S be a region of R^n. Then*

$$P((X_1, X_2, ..., X_n) \in S) = \underset{S}{\int \cdots \int} f(x_1, ..., x_n) \mathrm{d}x_1 ... \mathrm{d}x_n$$

3 Examples

1. Let X_1 and X_2 be discrete random variables with the following joint pdf:

$$p(x_1, x_2) = \begin{cases} \frac{p^{x_2}(1-p)^{1-x_2}}{x_2+1} & \text{for} x_1 = 0, x_2 \text{ and } x_2 = 0, 1 \\ 0 & \text{else} \end{cases}$$

Where $0 < p < 1$. Let's confirm that the joint pdf of X_1 and X_2 is valid. First, we note that the support of X_2 is $S_2 = \{0, 1\}$ and the support of X_1 is $S_1 = \{0, X_2\}$. This means that when $X_2 = 0$ the support of X_1 is just $\{0\}$, but when $X_2 = 1$, the support is $\{0, 1\}$. So, the support of one of the random vairables actually *depends* on the value of a different random variable

$$\begin{aligned}
\sum_{x_1 \in S_1} \sum_{x_2 \in S_2} p(x_1, x_2) &= \sum_{x_2 \in S_2} \sum_{x_1 \in S_1} p(x_1, x_2) \\
&\quad \uparrow \text{First, we change the order of the summation, since} \\
&\quad S_1 \text{ depends on the value of } x_2 \\
&= \sum_{x_2=0}^{1} \sum_{x_1=0}^{x_2} \frac{p^{x_2}(1-p)^{1-x_2}}{x_2+1} \\
&= \sum_{x_1=0}^{0} \frac{p^{(0)}(1-p)^{1-(0)}}{(0)+1} + \sum_{x_1=0}^{1} \frac{p^{(1)}(1-p)^{1-(1)}}{(1)+1} \\
&= \frac{(1-p)}{1} + \sum_{x_1=0}^{1} \frac{p}{2} \\
&= (1-p) + \left(\frac{p}{2} + \frac{p}{2}\right) \\
&= (1-p) + p \\
&= 1
\end{aligned}$$

2. Let X_1 and X_2 be continuous random variables with the following pdf:

$$f(x_1, x_2) = \begin{cases} exp(-x_1) & \text{for} 0 < x_1 \text{ and } 0 < x_2 < 1 \\ 0 & \text{else} \end{cases}$$

Lets find $P(.5 < X_1 + X_2 < 1)$. Note, we will apply the theorem we described earlier, where our region S is described as $.5 < X_1 + X_2 < 1$

$$\begin{aligned}
P(.5 < X_1 + X_2 < 1) &= \iint_{x_1+x_2=.5}^{1} f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \int_{x_2=0}^{.5}\int_{x_1=.5-x_2}^{1-x_2} f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 + \int_{x_2=.5}^{1}\int_{x_1=0}^{1-x_2} f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \int_{x_2=0}^{.5}\int_{x_1=.5-x_2}^{1-x_2} exp(-x_1)\mathrm{d}x_1\mathrm{d}x_2 + \int_{x_2=.5}^{1}\int_{x_1=0}^{1-x_2} exp(-x_1)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \int_{x_2=0}^{.5} [-exp(-x_1)]_{x_1=.5-x_2}^{1-x_2}\mathrm{d}x_2 + \int_{x_2=.5}^{1} [-exp(-x_1)]_{x_1=0}^{1-x_2}\mathrm{d}x_2 \\
&= \int_{x_2=0}^{.5} [exp(-(.5-x_2)) - exp(-(1-x_2))]\mathrm{d}x_2 \\
&\quad + \int_{x_2=.5}^{1} [exp(-(0)) - exp(-(1-x_2))]\mathrm{d}x_2 \\
&= \int_{x_2=0}^{.5} [exp(x_2-.5) - exp(x_2-1)]\mathrm{d}x_2 + \int_{x_2=.5}^{1} [1 - exp(x_2-1)]\mathrm{d}x_2 \\
&= [exp(x_2-.5) - exp(x_2-1)]_{x_2=0}^{.5} + [x_2 - exp(x_2-1)]_{x_2=.5}^{1} \\
&= [(exp((.5)-.5) - exp((.5)-1)) - (exp(0()-.5) - exp((0)-1))] \\
&\quad +[(1 - exp((1)-1)) - (.5 - exp((.5)-1))] \\
&= [1 - exp(-.5) + exp(-1) - exp(-.5)] + [exp(-.5) - .5] \\
&= .5 + exp(-1) - exp(-.5)
\end{aligned}$$

3. For this example we will simply introduce the bivariate normal distribution.

Definition 4. *Let X_1, X_2 be continuous random varables with the joint PDF*

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left(-\frac{1}{2(1-\rho^2)}((\frac{x_1-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2}) + (\frac{x_2-\mu_2}{\sigma_2})^2)\right)$$

Where $-\infty < \mu_1, \mu_2 < \infty$, $0 < \sigma_1, \sigma_2$, and $-1 < \rho < 1$. (Note, the supports of X_1 and X_2 are both $(-\infty, \infty)$). When X_1 and X_2 have th joint pdf with the parameters $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ asdescribed, then X_1 and X_2 are jointly said to have a bivariate normal distribution. This is denoted as $(X_1, X_2) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$