5.1: Jointly Distributed Random Variables

Taylor

University of Virginia

Motivation

When multiple rvs vary together, we need a **joint distribution** to fully describe them.

Definition

For discrete rvs, we have the **joint probability mass function**

$$p(x, y) = P(X = x \cap Y = y)$$

also, if A is a two-dimensional set:

$$P[(X, Y) \in A] = \sum_{(x,y) \in A} \sum p(x, y)$$

Definition

The **marginal probability mass functions** can be obtained from the joint via summation...

$$p_X(x) = \sum_y p(x, y)$$

and

$$p_Y(y) = \sum_x p(x, y)$$

Definition

Let X and Y be two cts rvs. Then $f(x,y)$ is the **joint probability density function** if for any A

$$P[(X,Y) \in A] = \iint_A f(x,y)dxdy$$

also:

1. $\iint_{\mathcal{S}} f(x,y)dxdy = 1$
2. $f(x,y) \geq 0$

A Word of Caution

We say

$$P[(X, Y) \in A] = \iint_A f(x, y)dxdy$$

...but doing this in practice is more difficult. We need to be very careful when we find the bounds of integration since we're dealing with more than one random variable.

Motivation

We can obtain marginal pdfs via integration

$$f_X(x) = \int_y f(x, y)dy$$

$$f_Y(y) = \int_x f(x, y)dx$$

Example 5.5 on page 237

Let $f(x, y) = 24xy$ with $0 < x < 1$, $0 < y < 1$ and $x + y \leq 1$. Verify that it's a pdf.

Example 5.5 on page 237

Let $f(x,y) = 24xy$ with $0 < x < 1$, $0 < y < 1$ and $x + y \leq 1$. Verify that it's a pdf.

First, it's obvious that this function can't be negative for any legitimate (x,y) couple. Next we have to make sure it integrates to 1.

$$\begin{aligned}
\iint_{\mathcal{S}} f(x,y)dxdy &= \int_0^1 \int_0^{1-x} 24xy \; dydx \\
&= \int_0^1 24x(y^2/2)|_{y=0}^{y=1-x} dx \\
&= 12 \int_0^1 x(1-x)^2 dx \\
&= 12 \int_0^1 x - 2x^2 + x^3 dx \\
&= 12 \left[\frac{x^2}{2} - \frac{2}{3}x^3 + \frac{1}{4}x^4 \right] \Big|_{x=0}^{x=1} = 1
\end{aligned}$$

Example 5.5 on page 237

Let $f(x, y) = 24xy$ with $0 < x < 1$, $0 < y < 1$ and $x + y \leq 1$. What's the probability that $X < Y$?

Example 5.5 on page 237

Let $f(x,y) = 24xy$ with $0 < x < 1$, $0 < y < 1$ and $x + y \leq 1$. What's the probability that $X < Y$?

Draw the picture for the bounds!

$$\begin{aligned} P(X < Y) &= \int_0^{.5} \int_x^{1-x} (24xy) dy dx \\ &= \int_0^{.5} 12xy^2 |_{y=x}^{y=1-x} dx \\ &= \int_0^{.5} 12x(1-x)^2 - 12x^3 dx \\ &= \int_0^{.5} 12x - 24x^2 dx \\ &= [6x^2 - 8x^3] |_{x=0}^{x=.5} = .5 \end{aligned}$$

Motivation

Earlier we said that a joint pdf/pmf completely describes rvs. A lot of times X and Y are **dependent**. If we knew something about, say X, then that influences what Y can be, and vice versa.

The best case scenario is when rvs are **independent**. It simplifies pretty much everything, and it's really big in statistics and probability.

Also: don't worry about it seeming unrealistic at the moment. Different types of independence assumptions are pretty much always used somewhere whenever anyone is doing any kind of modeling.

Definition

Two random variables X and Y are **independent** if, for any x, y pair

$$p(x, y) = p_X(x)p_Y(y)$$

if they're discrete, or

$$f(x, y) = f_X(x)f_Y(y)$$

when they're cts. If there exists at least one pair x, y such that this isn't true, then X and Y are **dependent**

Example

Example 5.8 on page 239: Let X and Y be two independent rvs that describe lifetimes of components of some machine. Let $X \sim \text{Exponential}(\lambda_1)$ and $Y \sim \text{Exponential}(\lambda_2)$. Write down the joint pdf of these two rvs. Then use that to find the probability of $A = \{X > 1500 \cap Y > 1500\}$

Example

Example 5.8 on page 239: Let X and Y be two independent rvs that describe lifetimes of components of some machine. Let $X \sim \text{Exponential}(\lambda_1)$ and $Y \sim \text{Exponential}(\lambda_2)$. Write down the joint pdf of these two rvs. Then use that to find the probability of $A = \{X > 1500 \cap Y > 1500\}$

$$f_{X,Y}(x,y) = \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 y} \qquad x, y > 0$$

$$\begin{aligned} P(A) &= \int_{1500}^{\infty} \int_{1500}^{\infty} \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 y} dx dy \\ &= \int_{1500}^{\infty} \lambda_1 e^{-\lambda_1 x} dx \int_{1500}^{\infty} \lambda_2 e^{-\lambda_2 y} dy \\ &= \exp\left[-\lambda_1 1500\right] \exp\left[-\lambda_2 1500\right] \end{aligned}$$

Definition

Let $X_1, \ldots, X_n$ be a collection of discrete rvs. Then the **joint pmf** is

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, \ldots, X_N = x_n)$$

Let $X_1, \ldots, X_n$ be a collection of cts rvs. Then the **joint pdf** is the function $f(x_1, \ldots x_n)$ such that for any n-dimensional rectangle

$$P(a_1 \leq X_1 \leq b_1, \ldots, a_n \leq X_n \leq b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) dx_n \cdots dx_1$$

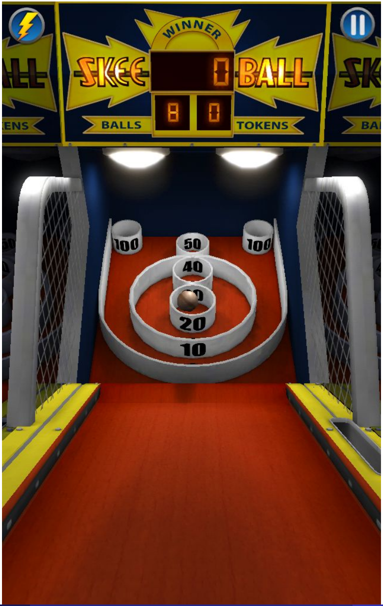
Motivation

We'll mention the multivariate generalization of the binomial distribution now. We won't prove much with it, but it's useful so it's good to have heard about it.

A good way to think about it is say you have $n = 10$ independent throws at a skee-ball machine. There are eight different holes, each with different point values. The higher the point value, the harder it is to get the ball in there.

To describe this you need a probability for each hole $p_1, \ldots, p_8$, and you need to know how many throws you get. Then you get a function that describes the random vector of where you throws ended up (e.g. $n = 10$, (5,4,0,0,0,0,1,0)). Notice how all those sum to $n = 10$, and the vector is eight dimensional.

Motivation



Definition

The multinomial distribution for the r-dimensional random count vector $(X_1, \ldots X_r)$ is

$$p(x_1, \ldots, x_r) = \frac{n!}{x_1! \cdots x_r!} p_1^{x_1} \cdots p_r^{x_r}$$

if $x_1, \ldots, x_r \in \{0, 1, \ldots, n\}$ with $\sum_i x_i = n$ and $\sum_i p_i = 1$

Note: set $r = 2$ and see how you get the binomial distribution

Example

Example 5.9 on page 241. Let X_1, X_2 and X_3 denote the number of AA, Aa and aa alleles, respectively, in $n = 10$ independent pea sections. Let $p_1 = p_3 = .25$ and $p_2 = .5$. Then we can find stuff like

$$P(X_1 = 2, X_2 = 5, X_3 = 3) = \frac{10!}{2!5!3!}(.25^2)(.5^5)(.25^3) = .0769$$

Definition

To define independence for larger collections of rvs, we can't just make each two-way pair independent. We have to talk about triples, quadruples, and so on.

The random variables $X_1, \ldots X_n$ are **mutually independent** if for every subset $X_{i_1}, \ldots X_{i_k}$

$$f_{X_{i_1},\ldots X_{i_k}}(x_{i_1}, \ldots x_{i_k}) = f_{X_{i_1}}(x_{i_1}) \cdots f_{X_{i_k}}(x_{i_k})$$

Definition

To define independence for larger collections of rvs, we can't just make each two-way pair independent. We have to talk about triples, quadruples, and so on.

The random variables $X_1, \ldots X_n$ are **mutually independent** if for every subset $X_{i_1}, \ldots X_{i_k}$

$$f_{X_{i_1},\ldots X_{i_k}}(x_{i_1}, \ldots x_{i_k}) = f_{X_{i_1}}(x_{i_1}) \cdots f_{X_{i_k}}(x_{i_k})$$

Note: a lot of times instead of writing $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ i'll just write $f_{\mathbf{X}}(\mathbf{x})$