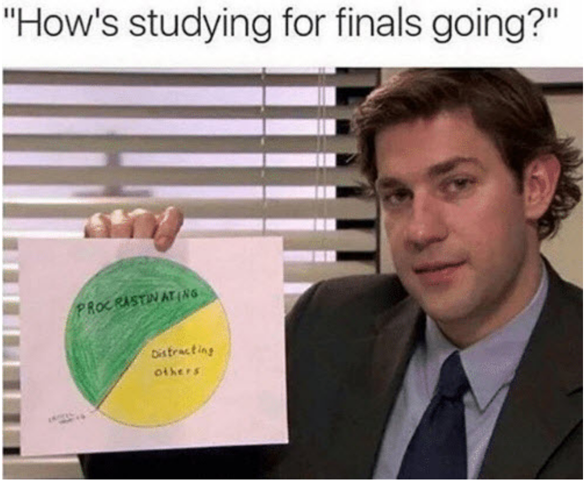
STAT 230 SECTION 2 LECTURE 30



Today's Agenda

Last time:

- Multivariate Distributions
  - joint probability function
  - marginal distributions

Today (Lec 29, 07/11):

- Independence
- Conditional distributions
- Multinomial Distribution

Definition

Suppose that X and Y are *discrete* random variables defined on the same sample space (in general, when we consider two or more random variables it is assumed they are defined on the same sample space.)

The **joint probability function** of X and Y is

$$f(x, y) = P(\{X = x\} \cap \{Y = y\}) \quad x \in X(S), y \in Y(S).$$

A shorthand for this is

$$f(x, y) = P(X = x, Y = y).$$

Properties of the joint probability function:

a) $f(x, y) \geq 0$

b) $\sum_{x,y} f(x, y) = 1.$

Point: joint probability function is still a probability function.

Definition

Suppose that X and Y are *discrete* random variables with joint probability function $f(x, y)$.
The **marginal probability function** of X is

$$f_X(x) = P(X = x) = \sum_{y \in Y(S)} f(x, y).$$

Similarly, the marginal distribution of Y is

$$f_Y(y) = P(Y = y) = \sum_{x \in X(S)} f(x, y).$$

Let's step it up a bit.

Example

Suppose X and Y have joint probability function

$$f(x,y) = \frac{1}{6}\left(\frac{1}{2}\right)^x \left(\frac{2}{3}\right)^y, \quad x,y = 0,1,2...$$

Compute the marginal probability functions $f_X(x)$ and $f_Y(y)$.

We make use of the identity

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \quad 0 < q < 1.$$

The margin of X is

$$\begin{aligned}
f_X(x) &= \sum_{y=0}^{\infty} \frac{1}{6} \left(\frac{1}{2}\right)^x \left(\frac{2}{3}\right)^y \\
&= \frac{1}{6} \left(\frac{1}{2}\right)^x \sum_{y=0}^{\infty} \left(\frac{2}{3}\right)^y \\
&= \frac{1}{6} \left(\frac{1}{2}\right)^x \frac{1}{1-\frac{2}{3}} \\
&= \frac{1}{2} \left(\frac{1}{2}\right)^x, \quad x = 0, 1, \ldots
\end{aligned}$$

from which we conclude that $X \sim Geo(1/2)$ (number of failures until first success).

Similarly, you find

$$f_Y(y) = \frac{2}{3} \left(\frac{2}{3}\right)^y, \quad y = 0, 1, \ldots$$

so that $Y \sim Geo(2/3)$.

Let's extend another existing concept to random variables.

Definition

X and Y are **independent** random variables if

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

for all values of (x, y).

In general, $X_1, X_2, \ldots, X_n$ are **independent** if

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \ldots f_n(x_n)$$

for all values of $(x_1, \ldots, x_n)$.

This means that if you find a single combination of $(x_1, \ldots, x_n)$ values that doesn't satisfy the above equation, then $X_1, \ldots, X_n$ are not independent.

Recall the following example:

Example

Suppose a fair coin is tossed 3 times. Define the random variables $X =$ "number of Heads", and

$$Y = \begin{cases} 1 & \text{Head occurs on the first toss,} \\ 0 & \text{Tail occurs on the first toss.} \end{cases}$$

Are X, Y independent? Justify.

Question

Suppose $X \sim Poi(2)$, $Y \sim Poi(3)$, and that X and Y are independent. What is the joint probability function of X and Y?

A $f(x, y) = e^{-6} \frac{2^x 3^y}{x!y!}, \;\; x, y = 0, 1, 2, ...$

B $f(x, y) = e^{-5} \frac{5^{x+y}}{x!y!}, \;\; x, y = 0, 1, 2, ...$

C $f(x, y) = e^{-5} \frac{2^x 3^y}{x!y!}, \;\; x, y = 0, 1, 2, ...$

D $f(x, y) = e^{-6} \frac{6^{x+y}}{x!y!}, \;\; x, y = 0, 1, 2, ...$

Conditional Distributions

For events A, B with $P(B) \neq 0$ we defined

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

We can now define conditional probability functions.

Definition

The **conditional probability function** of X given $Y = y$ is denoted $f_X(x|y)$, and is defined to be

$$f_X(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(x, y)}{f_Y(y)},$$

Given that $f_Y(y) > 0$. $f_Y(y|x)$ is similarly defined.

Let's revisit the coin flip example.

Example

Suppose a fair coin is tossed 3 times. Define the random variables $X =$ "number of Heads", and

$$Y = \begin{cases} 1 & \text{Head occurs on the first toss,} \\ 0 & \text{Tail occurs on the first toss.} \end{cases}$$

Find $f_X(x|y)$ for $y = 0, 1$.

With multiple random variables, we are even more interested in **functions** of such variables. For example, Let A, M, F be the random variables for your assignment, midterm and final grades respectively. Then, it's natural to consider the overall grade $G = g(A, M, F)$ as a function of random variables A, M, F.

In general, we have the following formula for the probability function of $U = g(X_1, X_2, \ldots, X_n)$.

$$P(U = u) = \sum_{\substack{(x_1,\ldots x_n) \text{ such that} \\ g(x_1,\ldots,x_n)=u}} f(x_1, \ldots, x_n)$$

Let's look at some useful results.

Theorem (Sum of independent Poisson is Poisson)

If $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$ independently, then $T = X + Y \sim Poi(\lambda_1 + \lambda_2)$.

Proof:

$$\begin{aligned}
P(X+Y=t) &= \sum_{(x,y):x+y=t} P(X=x, Y=y) \\
&= \sum_{x=0}^{t} P(X=x, Y=t-x) \\
&= \sum_{x=0}^{t} \underbrace{f(x, t-x)}_{=f_X(x)f_Y(t-x)} \\
&= \sum_{x=0}^{t} e^{-\lambda_1} \frac{\lambda_1^x}{x!} e^{-\lambda_2} \frac{\lambda_2^{t-x}}{(t-x)!} \ldots
\end{aligned}$$

. . . rest is exercise and in the course notes.

Another useful result proven very similarly:

Theorem (Sum of independent binomial is binomial)

If $X \sim Bin(n, p)$ *and* $Y \sim Bin(m, p)$ *independently, then* $T = X + Y \sim Bin(n + m, p)$

Note: This makes intuitive sense.

- X: number of successes in n independent trials with prob p
- Y: number of successes in m independent trials with prob p, independent of X
- $X + Y$: number of successes in $n + m$ independent trials with prob p.

Question

When dealing with a function of random variables, it's important to specify its range, just like the one variable case.

Suppose that $X \sim \mathit{Binomial}(10, 3)$ and $Y \sim \mathit{Poi}(2)$, and X and Y are independent. What is the range of the random variable U if $U = X - Y$?

A $U(S) = (-\infty, 0) \cap \mathbb{Z}$

B $U(S) = (-\infty, 10] \cap \mathbb{Z}$

C $U(S) = (-\infty, 0] \cap \mathbb{Z}$

D $U(S) = [0, \infty) \cap \mathbb{Z}$

E $U(S) = [0, 10] \cap \mathbb{Z}$

The sum of random variables can also "decompose" some distributions in terms of simpler ones.

Theorem

Let $X_1, X_2, \ldots, X_n$ each follow Bernoulli(p) independently. Then,

$$X_1 + X_2 + \ldots + X_n \sim Bin(n, p).$$

Similarly, we can specify the relationship between geometric and negative binomial distributions.

Theorem

Let $X_1, X_2, \ldots, X_k$ each follow $Geo(p)$ independently. Then,

$$X_1 + X_2 + \ldots + X_k \sim NB(k, p).$$

Theorem

Let $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$ independently. Then, given $X + Y = n$, X follows binomial distribution. That is,

$$X | X + Y = n \sim Bin\left(n, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right).$$

Similarly, for Y, we have

$$Y | X + Y = n \sim Bin\left(n, \frac{\lambda_2}{\lambda_1 + \lambda_2}\right)$$

Proof: Exercise. Use that $X + Y \sim Poi(\lambda_1 + \lambda_2)$.

Example

The duty don in residence receives two types of duty phone calls: emergency calls, and non-emergency calls. Emergency calls arrive according to Poisson distribution with $\lambda = 1$ per 6 hours. Non-emergency also arrive according to Poisson distribution with $\lambda = 3$ per 6 hours, independently of emergency calls.

- What is the distribution of the number of duty phone calls over 6 hours?
- What is the distribution of emergency calls over 6 hours, given that there are 10 calls in total over 6 hours?
- What is the distribution of non-emergency calls over 6 hours, given that there are 8 calls in total over 6 hours?