Parametric Statistics

Joint Distributions

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

October 9, 2023

Recap

- Random variables are functions from the sample space to the real line.
- Random variables can be discrete or continuous.
- Discrete RVs have probability mass functions.
- Continuous RVs have probability density functions.
- All RVs can be described with the cumulative distribution function.
- The quantile function is the inverse of the CDF, for continuous distributions.

Today

3.4 Bivariate Distributions
3.5 Marginal Distributions
3.6 Conditional Distributions
3.7 Multivariate Distributions
3.8 Functions of a Random Variable

Joint Probability Mass Function

y \ x	0	1	2	3
3	$1/20$	$2/20$	$1/20$	$1/20$
2	$2/20$	$4/20$	$1/20$	$1/20$
1		$2/20$	$4/20$	
0	$1/20$			

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

$$\sum_{all(x,y)} f(x, y) = 1 \text{ (still a probability mass function)}$$

Joint Probability Mass Function

y \ x	0	1	2	3
3	$1/20$	$2/20$	$1/20$	$1/20$
2	$2/20$	$4/20$	$1/20$	$1/20$
1		$2/20$	$4/20$	
0	$1/20$			

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

$$\sum_{all(x,y)} f(x, y) = 1 \text{ (still a probability mass function)}$$

What is $P(X \geq 2, Y \geq 2)$?

Marginal Probability Functions

y \ x	0	1	2	3
3	1/20	2/20	1/20	1/20
2	2/20	4/20	1/20	1/20
1		2/20	4/20	
0	1/20			

Marginal Distribution

If X and Y have a discrete joint distribution for which the joint p.f. is f, then the marginal p.f. f_1 of X is

$$f_1(x) = \sum_{\text{All } y} f(x, y).$$

Similarly, the marginal p.f. f_2 of Y is $f_2(y) = \sum_{\text{All } x} f(x, y)$.

Conditional Probability Mass Functions

y \ x	0	1	2	3
3	1/20	2/20	1/20	1/20
2	2/20	4/20	1/20	1/20
1		2/20	4/20	
0	1/20			

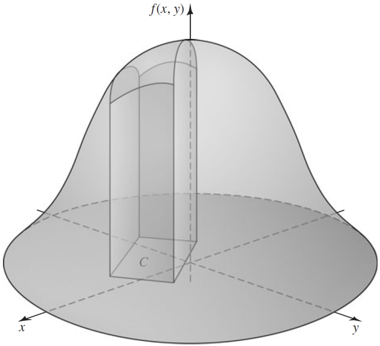
$y = 2$

Conditional Probability: $P(X = x|Y = y) = \dfrac{P(X = x, Y = y)}{P(Y = y)}$

$$\text{e.g., } P(x|y = 2) = \{2/8, 4/8, 1/8, 1/8\}$$

$$\sum_x P(x|y) = 1 \text{ (still a probability mass function)}$$

Joint Probability Density Function



Joint p.d.f.:

$$f(x, y) \geq 0, \text{everywhere}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

$$\Pr[(X, Y) \in C] = \int_C \int f(x, y) dx dy$$

Marginal PDFs

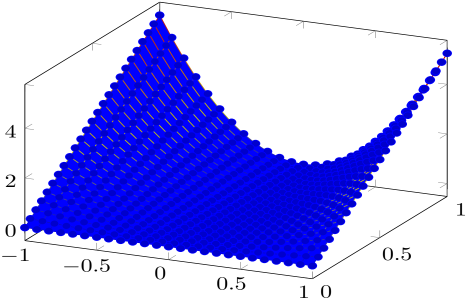
If X and Y have a continuous joint distribution with joint p.d.f. f, then the marginal p.d.f. f_1 of X is

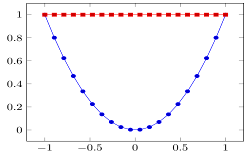
$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy \text{ for } -\infty < x < \infty.$$

Similarly, the marginal p.d.f. f_2 of Y is

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx \text{ for } -\infty < y < \infty.$$

Marginal PDFs





$$f(x,y) = \begin{cases} \frac{21}{4}x^2y & \text{for } x^2 \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find the marginal pdfs $f_1(x), f_2(y)$.

Conditional PDFs

Let X and Y have a continuous joint distribution with joint p.d.f. f and respective marginals f_1 and f_2. Let y be a value such that $f_2(y) > 0$. Then the conditional p.d.f. g_1 of X given that $Y = y$ is defined as follows:

$$g_1(x \mid y) = \frac{f(x,y)}{f_2(y)} \text{ for } -\infty < x < \infty.$$

For values of y such that $f_2(y) = 0$, we are free to define $g_1(x \mid y)$ however we wish, so long as $g_1(x \mid y)$ is a p.d.f. as a function of x.

Conditional PDFs

Example

A manufacturing process consists of two stages. The first stage takes Y minutes, and the whole process takes X minutes (which includes the first Y minutes). Suppose that X and Y have a joint continuous distribution with joint p.d.f.

$$f(x,y) = \begin{cases} e^{-x} & \text{for } 0 \le y \le x < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

After we learn how much time Y that the first stage takes, we want to update our distribution for the total time X. In other words, we would like to be able to compute a conditional distribution for X given $Y = y$. We cannot argue the same way as we did with discrete joint distributions, because $\{Y = y\}$ is an event with probability 0 for all y.

Conditional PDFs

Y is the time that the first stage of a process takes, while X is the total time of the two stages. We want to calculate the conditional p.d.f. of X given Y. We can calculate the marginal p.d.f. of Y as follows: For each y, the possible values of X are all $x \geq y$, so for each $y > 0$,

$$f_2(y) = \int_y^\infty e^{-x} dx =,$$

and $f_2(y) = 0$ for $y < 0$. For each $y \geq 0$, the conditional p.d.f. of X given $Y = y$ is then

$$g_1(x \mid y) = \frac{f(x, y)}{f_2(y)} =$$

Conditional PDFs

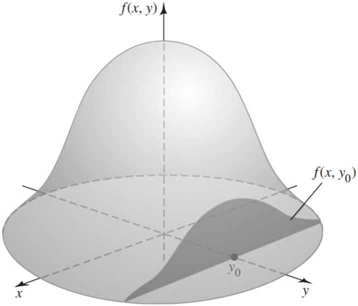
Y is the time that the first stage of a process takes, while X is the total time of the two stages. We want to calculate the conditional p.d.f. of X given Y. We can calculate the marginal p.d.f. of Y as follows: For each y, the possible values of X are all $x \geq y$, so for each $y > 0$,

$$f_2(y) = \int_y^\infty e^{-x} dx = e^{-y},$$

and $f_2(y) = 0$ for $y < 0$. For each $y \geq 0$, the conditional p.d.f. of X given $Y = y$ is then

$$g_1(x \mid y) = \frac{f(x, y)}{f_2(y)} = \frac{e^{-x}}{e^{-y}} = e^{y-x}, \text{ for } x \geq y$$

Conditional PDFs



Note: $g_1(x \mid y)$ is a limit:

$$g_1(x \mid y) = \lim_{\epsilon \to 0} \frac{\partial}{\partial x} \mathrm{P}(X \leq x \mid y - \epsilon < Y \leq y + \epsilon).$$

Joint/Conditional/Marginal PDFs

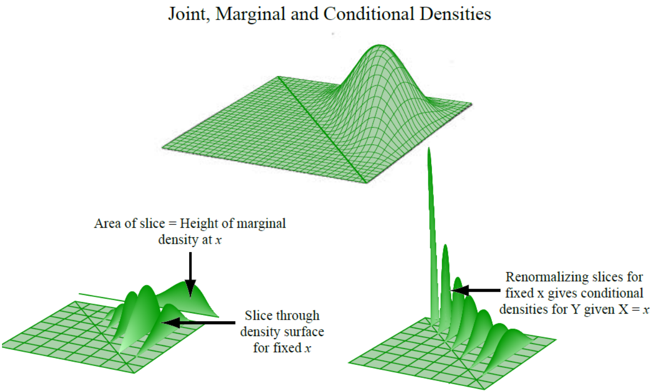


Image by MIT OpenCourseWare, adapted from *Probability*, by J. Pittman, 1999.

Independence of Random Variables

Definition (Independent Random Variables)

Two random variables are independent if for every two sets A and B in R the events $\{s : X(s) \in A\}$ and $\{s : Y(s) \in B\}$ are independent.

Theorem

Suppose that X and Y are random variables that have a joint p.f., p.d.f., or p.f./p.d.f. f. Two random variables X and Y are independent if and only if the following factorization is satisfied for all real numbers x and y :

$$f(x, y) = f_1(x) f_2(y).$$

Random Samples/i.i.d./Sample Size.

- Let f be a distribution.
- n random variables $X_1, \dots, X_n$ form a random sample from this distribution if these random variables are independent and the marginal p.f. or p.d.f. of each of them is f.
- AKA Independent and Identically Distributed (i.i.d.) random variables
- n is the sample size.
- $(x_1, x_2, \dots, x_n) \in R^n$:

$$g(x_1, \dots, x_n) = f(x_1) f(x_2) \cdots f(x_n).$$

Conditional Independence

y \ x	0	1	2	3
3	$1/20$	$2/20$	$2/20$	$1/20$
2	$2/20$	$4/20$	$1/20$	$1/20$
1		$2/20$	$3/20$	
0	$1/20$			

- Assume you know that $X < 2$ and $Y \geq 2$
- Are X and Y independent in this new universe?

Multiplication Rule for RVs.

$$f(x, y) = g_1(x \mid y) f_2(y) = f_1(x) g_2(y \mid x)$$

Law of Total Probability for RVs.

Discrete:

$$f_1(x) = \sum_y g_1(x \mid y) f_2(y),$$

Continuous:

$$f_1(x) = \int_{-\infty}^{\infty} g_1(x \mid y) f_2(y) dy.$$

Bayes Rule for RVs.

$$g_2(y \mid x) = \frac{g_1(x \mid y) f_2(y)}{f_1(x)}$$

Functions of Random Variables are Random Variables

Function of a Discrete Random Variable

$$g(y) = \Pr(Y = y) = \Pr[r(X) = y] = \sum_{x:r(x)=y} f(x).$$

Example

$$f(x) = \frac{x^2}{a}, \quad x \in \{-3, -2, -1, 1, 2, 3\}$$

- Find a.
- Find the pmf of X
- Find the pmf of $Z = X^2$

Mixed Distributions

Joint p.f./p.d.f

Let X and Y be random variables such that X is discrete and Y is continuous. Suppose that there is a function $f(x, y)$ defined on the xy-plane such that, for every pair A and B of subsets of the real numbers,

$$\Pr(X \in A \text{ and } Y \in B) = \int_B \sum_{x \in A} f(x, y) dy,$$

if the integral exists. Then the function f is called the joint p.f./p.d.f. of X and Y.

Joint (Cumulative) Distribution Function/c.d.f.

The joint distribution function or joint cumulative distribution function (joint c.d.f.) of two random variables X and Y is defined as the function F such that for all values of x and $y(-\infty < x < \infty$ and $-\infty < y < \infty)$

$$F(x, y) = \Pr(X \leq x \text{ and } Y \leq y).$$

Practice

A fair coin is tossed three times. Let

- X: number of heads on the first toss
- Y: total number of heads
- Find the joint distribution of X, Y.
- Find the marginal distributions of X and Y.
- Find the conditional distribution of $Y|X$
- Are X and Y independent?

Recap

- Two random variables have a bivariate joint distribution.
- More than two RVs have a multivariate joint distribution.
- We can compute marginal, conditional distributions from the joint pf.
- Independence is defined for RVs.
- Functions of RVs are RVs.