Parametric Statistics

Random Variables

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

October 4, 2023

Lecture Summary

3.1 Discrete Random Variables

3.2 Continuous Random Variables

3.3 The Cumulative Distribution Function

3.4 Bivariate Distributions

3.5 Marginal Distributions

Some slides from MIT open courseware

Last time

- Two events are called independent when knowing the value of one doesn't influence the probability of the value of the other.
- The conditional probability of A given B denotes the probability of event A in a world where B has occurred.
- We can use the multiplication to compute conditional, marginal and joint probabilities.
- Bayes rule connects $P(A|B)$ and $P(B|A)$. These two are confused but they are not the same.

Why is Bayes Rule so important?

- V: vaccinated
- H: hospitalized.
- $P(H|V) = 0.01$
- $P(H|V^c) = 0.2$
- Three different possibilities: $P(V) = 0.8, 0.5, 0.95$

Let' s use Bayes rule to compute $P(V|H)$ for all three cases.

$P(V)$	$P(H|V)$	$P(H|V^c)$	$P(V|H)$
0.5	0.01	0.2	0.0476
0.8	0.01	0.2	0.16667
0.95	0.01	0.2	0.4872

Random Variables

Random Variable

A random variable is a mapping $X : \Omega \to \mathbb{R}$ that assigns a real number $X(\omega)$ to each outcome ω.

Example

Consider the experiment of flipping a coin 10 times, and let $X(\omega)$ denote the number of heads in the outcome ω. For example, if $\omega = HHHTHHTTHT$, $X(\omega) = 6$.

Random Variables

- Why do we need random variables? (easier to work with than original sample space).
- A random variable is NOT a variable (in the algebraic sense).
- A random variable takes a specific value AFTER the experiment is conducted.

Notation

- Letter near the end of the alphabet since it is a variable in the context of the experiment.
- Capital letter to distinguish from algebraic variable.
- Lower case denotes a specific value of the random variable.

Random Variable: Definition

- An assignment of a value (number) to every possible outcome.
- A real-valued function of the sample space:

$$f : \Omega \to \mathbb{R}$$

- Can take discrete or continuous values.
- Discrete RVs have probability mass functions.
- Continuous RVs have probability density functions.

Discrete Random Variables

Discrete Distribution/Random Variable

We say that a random variable X has a discrete distribution or that X is a discrete random variable if X can take only a finite number k of different values $x_1, \ldots, x_k$ or, at most, an infinite sequence of different values $x_1, x_2, \ldots$.

Probability Function/p.f./Support.

If a random variable X has a discrete distribution, the probability function (abbreviated p.f.) of X is defined as the function f such that for every real number x,

$$f(x) = \Pr(X = x).$$

The closure of the set $\{x : f(x) > 0\}$ is called the support of (the distribution of) X.

Discrete Random Variables

Discrete Distribution/Random Variable

We say that a random variable X has a discrete distribution or that X is a discrete random variable if X can take only a finite number k of different values $x_1, \ldots, x_k$ or, at most, an infinite sequence of different values $x_1, x_2, \ldots$.

Probability Function/p.f./Support.

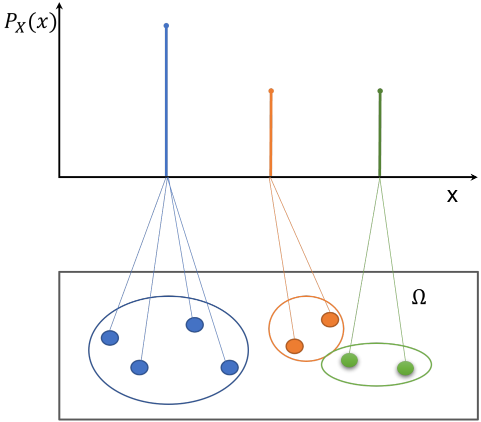
If a random variable X has a discrete distribution, the probability function (abbreviated p.f.) of X is defined as the function f such that for every real number x,

$$f(x) = \Pr(X = x).$$

The closure of the set $\{x : f(x) > 0\}$ is called the support of (the distribution of) X.

- Also known as the probability mass function.

Probability (mass) function



If $C \subset \mathbb{R}$:

$$\Pr(X \in C) = \sum_{x_i \in C} f(x_i),$$

If $x_1, x_2, \ldots$ includes all the possible values of X, then

$$\sum_{i=1}^{\infty} f(x_i) = 1$$

Bernoulli Distribution

Example: Coin Toss

- You toss a biased coin.
- $P(Heads) = p$
- $\Omega = \{Heads, Tails\}$
- Define $X(Heads) = 1$, $X(Tails) = 0$
- $f(x) =$

Bernoulli Distribution

Example: Coin Toss

- You toss a biased coin.
- $P(Heads) = p$
- $\Omega = \{Heads, Tails\}$
- Define $X(Heads) = 1$, $X(Tails) = 0$
- $f(x) = p$ for $x = 1$

Bernoulli Distribution

Example: Coin Toss

- You toss a biased coin.
- $P(Heads) = p$
- $\Omega = \{Heads, Tails\}$
- Define $X(Heads) = 1$, $X(Tails) = 0$
- $f(x) = p$ for $x = 1$
  $f(x) =$

Bernoulli Distribution

Example: Coin Toss

- You toss a biased coin.
- $P(Heads) = p$
- $\Omega = \{Heads, Tails\}$
- Define $X(Heads) = 1$, $X(Tails) = 0$
- $f(x) = p$ for $x = 1$
  $f(x) = 1 - p$ for $x = 0$.

Bernoulli Distribution

Example: Coin Toss

- You toss a biased coin.
- $P(Heads) = p$
- $\Omega = \{Heads, Tails\}$
- Define $X(Heads) = 1$, $X(Tails) = 0$
- $f(x) = p$ for $x = 1$
  $f(x) = 1 - p$ for $x = 0$.

Distribution

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{for } x = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

Discrete Uniform Distribution

Example: Lottery

- Select a three digit number (leading 0 s allowed).
- Each digit is chosen at random from a separate urn with 10 balls, each with a different digit.
- $\Omega = (i_1, i_2, i_3)$ where $i_j \in \{0, \dots, 9\}$ for $j = 1, 2, 3$
- If $\omega = (i_1, i_2, i_3)$, define $X(\omega) = 100i_1 + 10i_2 + i_3$. For example, $X(0, 1, 5) = 15$
- $f(x) =$

Discrete Uniform Distribution

Example: Lottery

- Select a three digit number (leading 0 s allowed).
- Each digit is chosen at random from a separate urn with 10 balls, each with a different digit.
- $\Omega = (i_1, i_2, i_3)$ where $i_j \in \{0, \ldots, 9\}$ for $j = 1, 2, 3$
- If $\omega = (i_1, i_2, i_3)$, define $X(\omega) = 100i_1 + 10i_2 + i_3$. For example, $X(0, 1, 5) = 15$
- $f(x) = 0.001$ for each integer $x \in \{0, 1, \ldots, 999\}$.

Discrete Uniform Distribution

Example: Lottery

- Select a three digit number (leading 0 s allowed).
- Each digit is chosen at random from a separate urn with 10 balls, each with a different digit.
- $\Omega = (i_1, i_2, i_3)$ where $i_j \in \{0, \ldots, 9\}$ for $j = 1, 2, 3$
- If $\omega = (i_1, i_2, i_3)$, define $X(\omega) = 100i_1 + 10i_2 + i_3$. For example, $X(0, 1, 5) = 15$
- $f(x) = 0.001$ for each integer $x \in \{0, 1, \ldots, 999\}$.

Distribution

$$f(x) = \begin{cases} \frac{1}{b-a+1} & \text{for } x = a, \ldots, b, \\ 0 & \text{otherwise.} \end{cases}$$

Continuous Random Variables

Continuous Distribution/Random Variable

We say that a random variable X has a continuous distribution or that X is a continuous random variable if there exists a non-negative function f, defined on the real line, such that for every interval of real numbers (bounded or unbounded), the probability that X takes a value in the interval is the integral of f over the interval.

Probability Function/p.f./Support.

If X has a continuous distribution, the function f is called the probability density function (abbreviated p.d.f.) of X. The closure of the set $\{x : f(x) > 0\}$ is called the support of (the distribution of) X.

Continuous Random Variables

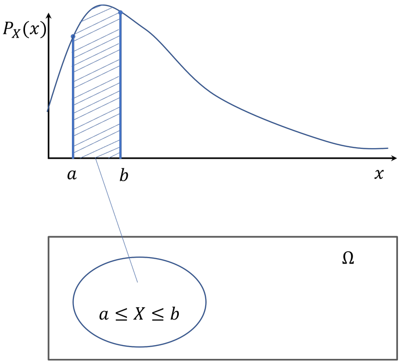
Continuous Distribution/Random Variable

We say that a random variable X has a continuous distribution or that X is a continuous random variable if there exists a non-negative function f, defined on the real line, such that for every interval of real numbers (bounded or unbounded), the probability that X takes a value in the interval is the integral of f over the interval.

Probability Function/p.f./Support.

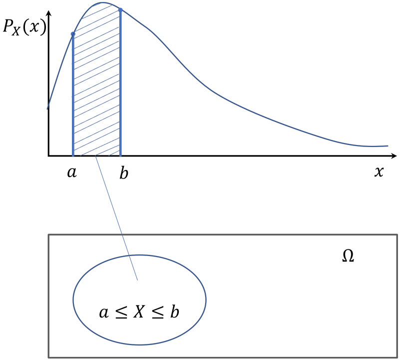
If X has a continuous distribution, the function f is called the probability density function (abbreviated p.d.f.) of X. The closure of the set $\{x : f(x) > 0\}$ is called the support of (the distribution of) X.

Continuous Random Variables



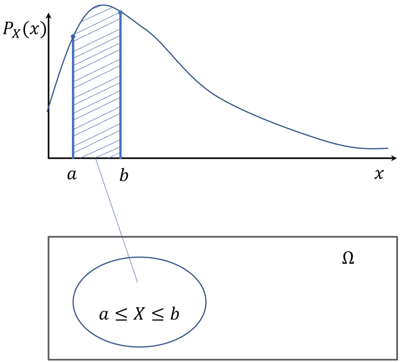
- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$

Continuous Random Variables



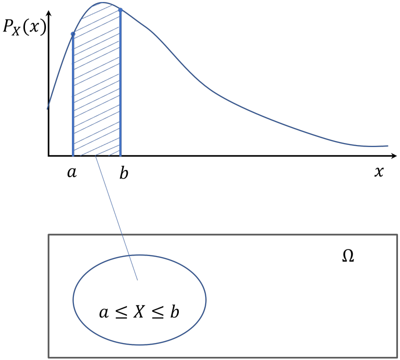
- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$
- $P(a \leq X) = \int_a^\infty f_X(x)dx$

Continuous Random Variables



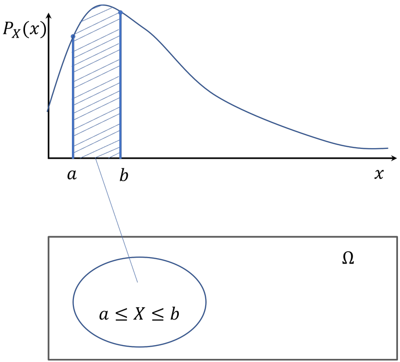
- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$
- $P(a \leq X) = \int_a^\infty f_X(x)dx$
- $P(X = a) = \int_a^a f_X(x)dx = 0$

Continuous Random Variables



- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$
- $P(a \leq X) = \int_a^\infty f_X(x)dx$
- $P(X = a) = \int_a^a f_X(x)dx = 0$
- $\int_{-\infty}^{\infty} f_X(x) = 1$

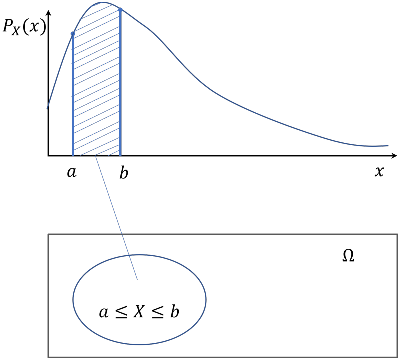
Continuous Random Variables



- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$
- $P(a \leq X) = \int_a^\infty f_X(x)dx$
- $P(X = a) = \int_a^a f_X(x)dx = 0$
- $\int_{-\infty}^\infty f_X(x) = 1$

Density is not probability!

Continuous Random Variables



- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$
- $P(a \leq X) = \int_a^\infty f_X(x)dx$
- $P(X = a) = \int_a^a f_X(x)dx = 0$
- $\int_{-\infty}^\infty f_X(x) = 1$

Density is not probability!

$$P(a \leq X \leq a + \epsilon = \int_a^{a+\epsilon} f_X(x)dx = f(a)\epsilon$$

$$P(b \leq X \leq b + \epsilon = \int_b^{b+\epsilon} f_X(x)dx = f(b)\epsilon$$

Continuous Uniform Distribution

Example: Weather

- Temperature Forecasts announce high and low temperature forecasts as integer numbers of degrees.
- In reality, the degrees are rounded up to the closest integer.
- Suppose that the forecaster announces a high temperature of y.
- X: the actual high temperature.
- X was equally likely to be any number in the interval from $y - 1/2$ to $y + 1/2$.

Continuous Uniform Distribution

Example: Weather

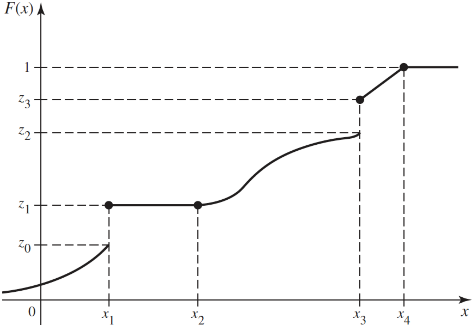
- Temperature Forecasts announce high and low temperature forecasts as integer numbers of degrees.
- In reality, the degrees are rounded up to the closest integer.
- Suppose that the forecaster announces a high temperature of y.
- X: the actual high temperature.
- X was equally likely to be any number in the interval from $y - 1/2$ to $y + 1/2$.

Distribution on the interval $[a, b]$

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{otherwise} \end{cases}$$

Cumulative Distribution Function (CDF)

$$F_X(x) = P(X \leq x) = \begin{cases} \int_{-\infty}^{x} f_X(t)dt, & \text{if } X \text{ is continuous} \\ \sum_{k \leq x} p_X(x) & \text{if } X \text{ is discrete} \end{cases}$$



Properties of the CDF

- Nondecreasing. The function $F(x)$ is nondecreasing as x increases; that is, if $x_1 < x_2$, then $F(x_1) \leq F(x_2)$.

Properties of the CDF

- Nondecreasing. The function $F(x)$ is nondecreasing as x increases; that is, if $x_1 < x_2$, then $F(x_1) \leq F(x_2)$.
- Limits at $\pm\infty$. $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

Properties of the CDF

- Nondecreasing. The function $F(x)$ is nondecreasing as x increases; that is, if $x_1 < x_2$, then $F(x_1) \leq F(x_2)$.
- Limits at $\pm\infty$. $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.
- Continuity from the Right. A c.d.f. is always continuous from the right; that is, $F(x) = F(x^+)$ at every point x.

Properties of the CDF

- For every value x,

$$\Pr(X > x) = 1 - F(x).$$

Properties of the CDF

- For every value x,

$$\Pr(X > x) = 1 - F(x).$$

- For all values x_1 and x_2 such that $x_1 < x_2$,

$$\Pr\left(x_1 < X \leq x_2\right) = F\left(x_2\right) - F\left(x_1\right).$$

Properties of the CDF

- For every value x,

$$\Pr(X > x) = 1 - F(x).$$

- For all values x_1 and x_2 such that $x_1 < x_2$,

$$\Pr\left(x_1 < X \leq x_2\right) = F\left(x_2\right) - F\left(x_1\right).$$

- **Proof.** Let $A = \left\{x_1 < X \leq x_2\right\}, B = \{X \leq x_1\}$, and $C = \{X \leq x_2\}$. A and B are disjoint, and their union is C, so

$$\Pr\left(x_1 < X \leq x_2\right) + \Pr\left(X \leq x_1\right) = \Pr\left(X \leq x_2\right).$$

Properties of the CDF

Let X have a continuous distribution, and let $f(x)$ and $F(x)$ denote its p.d.f. and the c.d.f., respectively. Then F is continuous at every x,

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

and

$$\frac{dF(x)}{dx} = f(x),$$

at all x such that f is continuous.

Quantiles/Percentiles

Fair Bets

Suppose that X is the number of points of the winning team on NBA basketball game, and X has c.d.f. F. Suppose that we want to place an even-money bet on X as follows: If $X \leq x_0$, we win one dollar and if $X > x_0$ we lose one dollar.
How can you make this bet fair?

Quantiles/Percentiles

Fair Bets

Suppose that X is the number of points of the winning team on NBA basketball game, and X has c.d.f. F. Suppose that we want to place an even-money bet on X as follows: If $X \leq x_0$, we win one dollar and if $X > x_0$ we lose one dollar.
How can you make this bet fair? $\Pr(X \leq x_0) = \Pr(X > x_0) = 1/2$.

Quantiles/Percentiles.

Let X be a random variable with c.d.f. F. For each p strictly between 0 and 1 , define $F^{-1}(p)$ to be the smallest value x such that $F(x) \geq p$. Then $F^{-1}(p)$ is called the p quantile of X or the $100p$ percentile of X. The function F^{-1} defined here on the open interval $(0, 1)$ is called the quantile function of X.

Quantiles/Percentiles

Fair Bets

Suppose that X is the number of points of the winning team on NBA basketball game, and X has c.d.f. F. Suppose that we want to place an even-money bet on X as follows: If $X \leq x_0$, we win one dollar and if $X > x_0$ we lose one dollar.
How can you make this bet fair? $\Pr(X \leq x_0) = \Pr(X > x_0) = 1/2$.

Quantiles/Percentiles.

Let X be a random variable with c.d.f. F. For each p strictly between 0 and 1 , define $F^{-1}(p)$ to be the smallest value x such that $F(x) \geq p$. Then $F^{-1}(p)$ is called the p quantile of X or the $100p$ percentile of X. The function F^{-1} defined here on the open interval $(0, 1)$ is called the quantile function of X.

- Example: Compute the quantile function for the uniform distribution.

Joint Probability Mass Function

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

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

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

$$\text{Marginal Probability:} f(x) = \sum_{y} f(x, y) \text{ (sum over all possible y)}$$

Joint Probability Mass Function

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

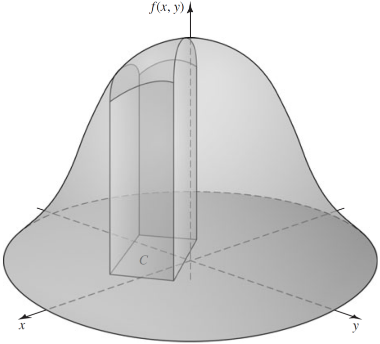
$$f(x, y) = P(X = x, Y = y)$$

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

Marginal Probability: $f(x) = \sum_{y} f(x, y)$ (sum over all possible y)

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

Joint Probability Density Function



Joint p.d.f.:

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

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

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

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 $-\infty < y < \infty)$

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

Recap

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