CS6462

Probabilistic and Explainable AI

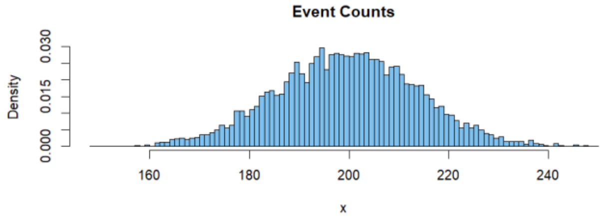
Lesson 5

Random Variables and Probability Distribution

by Emil Vassev

January 24, 2025

Random Variables and Probability



Definition: Random variables are represented by:

- a function that can take on either a finite number of values, each with an associated probability, or an infinite number of values, whose probabilities are summarized by a density function
- a function that assigns a numerical value to each outcome in ***S***, i.e., a real-valued function defined on ***S***

Example:

- If a coin is tossed three times, the sample space might be described by a list of 8 three-letter words,

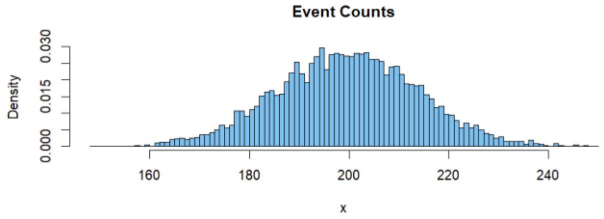
  S = {TTT, TTH, THT, HTT, HHT, HTH, THH, HHH}

- possible random variables:

  $\boldsymbol{F_X(S) = (\#H\text{'s in the word}) = \{0, 1, 2, 3\}}$, $\boldsymbol{X}$= #H's in the word

  $\boldsymbol{F_Y(S) = (\#T\text{'s in the word}) = \{0, 1, 2, 3\}}$, $\boldsymbol{Y}$ = #T's in the word

Discrete Random Variable



Definition:

- a random variable ***X*** is said to be discrete if it can assume only a finite or countable infinite number of distinct values
- a discrete random variable can be defined on both a countable or uncountable sample space

Example:

***S* = {TTT, TTH, THT, HTT, HHT, HTH, THH, HHH}**
***X* = #**H's in the word, ***X*= {0, 1, 2, 3}**

X=2, {HHT,HTH,THH}
P(x) = 3/8

Probability:

- the probability that ***X*** takes on the value ***x***, ***P(X=x)***, is defined as the sum of the probabilities of all sample points in ***S*** that are assigned the value ***x***
- ***P(X=x) = p(x)*** **-** function that assigns probabilities to each possible value x

p(x) **-** the probability function for ***X***

Probability Distribution

Definition:

- The probability distribution for a random variable describes how the probabilities are distributed over the values of the random variable.
- For a discrete random variable $\boldsymbol{X}$ the probability distribution is defined by a *probability mass function*, denoted by $\boldsymbol{F_X(X)}$.
- $\boldsymbol{F_X(x)}$ - provides the probability for each value of the random variable

  $\boldsymbol{F_X(x) = p(x) = P(X=x)}$ for each $\boldsymbol{x}$ within the range of $\boldsymbol{X}$

Rules for mass function $\boldsymbol{F_X(x)}$ for a discrete random variable $\boldsymbol{X}$:

(1) $\boldsymbol{F_X(x) \geq 0}$ - must be nonnegative for each value of the random variable $\boldsymbol{X}$

(2) $\boldsymbol{\sum_x F_X(x) = 1}$, sum of probabilities for each value of the random variable must = 1

Probability Distribution - Example

$P\,(\text{heads}) = \frac{1}{2} = 0.5$

Example:

Find a formula for the probability distribution of the total number of heads obtained in four tosses of a balanced coin.

S = see table

X= #H’s in the word

P(X=x):

- ***P(X = 0) = 1/16***
- ***P(X = 1) = 4/16***
- ***P(X = 2) = 6/16***
- ***P(X = 3) = 4/16***
- ***P(X = 4) = 1/16***

sample space, probabilities and random variable

Element of sample space	Probability	Value of random variable X (x)
HHHH	1/16	4
HHHT	1/16	3
HHTH	1/16	3
HTHH	1/16	3
THHH	1/16	3
HHTT	1/16	2
HTHT	1/16	2
HTTH	1/16	2
THHT	1/16	2
THTH	1/16	2
TTHH	1/16	2
HTTT	1/16	1
THTT	1/16	1
TTHT	1/16	1
TTTH	1/16	1
TTTT	1/16	0

Probability Distribution – Example (cont.)

$P(\text{heads}) = \frac{1}{2} = 0.5$

$$P(X=0) = \frac{1}{16},\ P(X=1) = \frac{4}{16},\ P(X=2) = \frac{6}{16},\ P(X=3) = \frac{4}{16},\ P(X=4) = \frac{1}{16}$$

- denominators of the five fractions are the same and the numerators of the five fractions are 1,4,6,4,1
- the numbers in the numerators form a set of ***binomial coefficients*** $\binom{n}{k} = \frac{n!}{k!\cdot(n-k)!}$

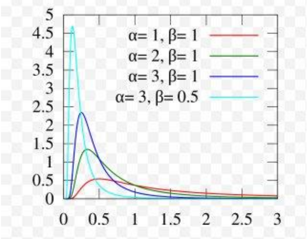
n = 4, k = x

$$\frac{1}{16} = \binom{4}{0} \quad \frac{4}{16} = \binom{4}{1} \quad \frac{6}{16} = \binom{4}{2} \quad \frac{4}{16} = \binom{4}{3} \quad \frac{1}{16} = \binom{4}{4}$$

- ***$F_X(x)$*** *– the probability mass function can be written as:*

$$f(x) = \frac{\binom{4}{x}}{16}\ for\ x = 0,\ 1,\ 2,\ 3,\ 4$$

Cumulative Distribution Function



*Definition: cumulative distribution function of **X***

If ***X*** is a discrete random variable, the function given by:

value of interest

$F_X(x) = P(x \le X) = \Sigma_{t \le x} f(t)$, for all $t \in R$, $t \le x$, $x \in X$

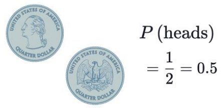
$X = \{x_1, x_2, x_3, x_4, x_5)$

where ***f(t)*** is the value of the probability distribution of ***X*** at ***t***

Rules for cumulative distribution function $F_X(x)$ of a discrete random variable ***X***:

1) $f(-\infty) = 0$ ***and*** $f(\infty) = 1$

2) if $a < b$, then $f(a) \le f(b)$ for any real numbers ***a*** and ***b*** within the range of ***X***



Cumulative Distribution Function - Example

Example: tossing a coin four times

determine the cumulative distribution function

$F_X(x) = P(x \leq X) = \Sigma_{t \leq x}\ f(t)$, for all ***$t \in R$***

- cumulative distribution function:

$$F(0) = f(0) = \frac{1}{16}$$

$$F(1) = f(0) + f(1) = \frac{1}{16} + \frac{4}{16} = \frac{5}{16}$$

$$F(2) = f(0) + f(1) + f(2) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16}$$

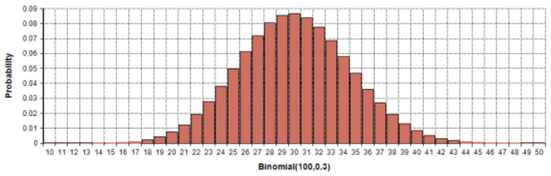
$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{6} = \frac{15}{16}$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{6} + \frac{1}{16} = \frac{16}{16}$$

$$F(x) = \begin{cases} 0 & for\ x < 0 \\ \frac{1}{16} & for\ 0 \leq x < 1 \\ \frac{5}{16} & for\ 1 \leq x < 2 \\ \frac{11}{16} & for\ 2 \leq x < 3 \\ \frac{15}{16} & for\ 3 \leq x < 4 \\ 1 & for\ x \geq 4 \end{cases}$$

sample space, probabilities and random variable

Element of sample space	Probability	Value of random variable X (x)
HHHH	1/16	4
HHHT	1/16	3
HHTH	1/16	3
HTHH	1/16	3
THHH	1/16	3
HHTT	1/16	2
HTHT	1/16	2
HTTH	1/16	2
THHT	1/16	2
THTH	1/16	2
TTHH	1/16	2
HTTT	1/16	1
THTT	1/16	1
TTHT	1/16	1
TTTH	1/16	1
TTTT	1/16	0



Binomial Distribution

Definition:

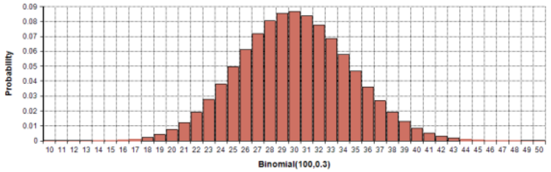
- a finite discrete distribution of the discrete random variable ***X***
- arises in situations where a sequence of what is known as *Bernoulli trials* is observed
- a *Bernoulli trial* is an experiment which has exactly two possible outcomes: *success* and *failure*
- *probability of success* is a fixed number **θ** which does not change with the number of experiments ***n***

The probability mass function for a binomial distribution:

$$b(x; n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n$$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Binomial Distribution - Example



Example: tossing a coin twice

What is the probability of getting one or more heads?

X= #H's in the word

The probability mass function for a binomial distribution:

The binomial distribution consists of the probabilities of each of the possible numbers of successes on ***n*** trials for independent events that each have a probability of **θ** of occurring.

$$b(x; n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n$$

For the coin flip example, n = 2:

***X=0, θ = ¼ , b(x; n, θ) =* 0.25**

***X=1, θ = ½ , b(x; n, θ) =* 0.5**

***X=2, θ = ½ , b(x; n, θ) =* 0.25**

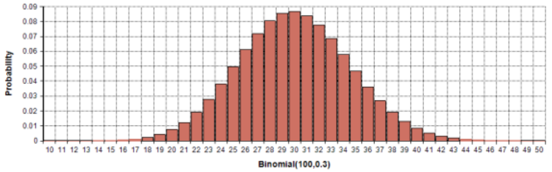
sample space

Outcome	First Flip	Second Flip
1	Heads	Heads
2	Heads	Tails
3	Tails	Heads
4	Tails	Tails

probabilities

Number of Heads	Probability
0	1/4
1	1/2
2	1/4

The probability of getting one or more heads is **0.5 + 0.25 = 0.75**



Poisson Distribution

Definition:

- used to model a number of events ***x*** occurring within a given time interval of an experiment.

- the formula for the Poisson *probability mass function* is:

$$\boldsymbol{P\,(x;\lambda) = \frac{e^{-\lambda} * \lambda^{x}}{x!}}$$

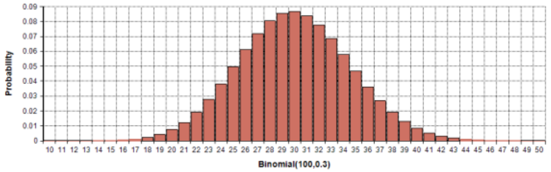
, for ***x*** = 0, 1, 2, ⋯ ; ***λ*** does not have to be an integer

λ is a shape parameter that indicates the average number of events in the given time interval

- *Poisson cumulative probability distribution function:*

$$\boldsymbol{F\,(x;\lambda) = \sum_{i}^{x} \frac{e^{-\lambda} * \lambda^{i}}{i!}}$$

Poisson Distribution - Example



Example:

Radioactivity example with an average of 2 decays/sec.

a) What's the probability of zero decays in one second?

b) What's the probability of more than one decay in one second?

Radioactive decay is the process by which an unstable atomic nucleus loses energy by radiation.

$$P(x;\lambda) = \frac{e^{-\lambda} * \lambda^x}{x!}$$

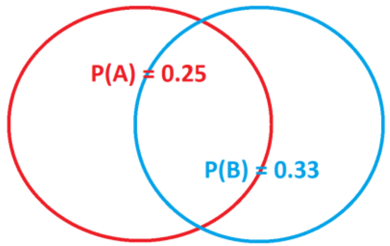
Solution a):

$$p(0,2) = \frac{e^{-2}2^0}{0!} = \frac{e^{-2}\cdot 1}{1} = e^{-2} = 0.135 \rightarrow 13.5\%$$

Solution b):

$$p(>1,2) = 1 - p(0,2) - p(1,2) = 1 - \frac{e^{-2}2^0}{0!} - \frac{e^{-2}2^1}{1!} = 1 - e^{-2} - 2e^{-2} = 0.594 \rightarrow 59.4\%$$

Summary



Random Variable:

- a function that assigns a numerical value to each outcome in ***S***
- discrete random variable

Probability Distribution for Random Variables

- probability mass function
- cumulative distribution function
- Binomial distribution
- Poisson distribution

Next Lesson – Bayes' Theorem

Thank You!

Questions?