Session 2: Probability distributions and density functions

Susan Thomas

http://www.igidr.ac.in/~susant

susant@mayin.org

IGIDR

Bombay

Session 2: Probability distributions and density functions - p. 1

Recap

The definition and scope of probability in the domain of an event space.

- Basic properties of probability
- Notion of mutually exclusive
- Notion of conditional and unconditional probability
- Notion of independance

The aim of this session

- 1. Discrete and continuous random variables
- 2. Probability distributions
- 3. Case 1: The bernoulli distribution
- 4. Case 2: The binomial distribution
- 5. Probability density functions
- 6. Case 3: The uniform distribution
- 7. Case 4: The normal distribution
- 8. Cumulative distributions and density functions
- 9. Sampling from distributions: population vs. sample

Random variables

Random Variables

A random variable is what an outcome can be. It could also be a function of an outcome.
 For example, the value of the first roll of a die.
 For example, the sum of a roll of two dice.

- Discrete random variables: where the possible events are countable.
 - For example, the roll of a dice, or the outcome of a horse race, or whether the firm will default or not.
- Continuous random variables: where the possible events are not countable.
 - For example, the number of white hair on my head, or how much dividend INFOSYSTCH will announce next year, or the price of Citibank stock.
- **Note:** Continuous RVs can have a fixed minimum of

Probability distributions

What is a probability distribution?

For a discrete RV, the probability distribution (PD) is a table of all the events and their related probabilities.

For example, in the roll of a die:

Value	Probability	Value	Probability
1	1/6	4	1/6
2	1/6	5	1/6
3	1/6	6	1/6

A probability distribution will contain all the outcomes and their related probabilities, and the probabilities will sum to 1.

How to read a probability distribution

From the distribution, we can find: X = 3 $\Pr(X = 3) = 1/6$ $\mathbf{X} = even number$ Pr(X = 2 or X = 4 or X = 6) = 3/6 = 1/2• More interesting, we can also find: Pr(X > 3) = 3/6 = 1/2This is called a cumulative probability.

tribution (CD)?

A table of the probabilities cumulated over the events.

	Value	Probability	Value	Probability
•	$X \leq 1$	1/6	$X \leq 4$	4/6
	$X \leq 2$	2/6	$X \leq 5$	5/6
	$X \leq 3$	3/6	X _{≤6}	1
-	Value $X \le 1$ $X \le 2$ $X \le 3$	Probability 1/6 2/6 3/6	Value $X \le 4$ $X \le 5$ $X \le 6$	Probability 4/6 5/6 1

The CD is a monotonically increasing set of numbers

The CD always ends with at the highest value of 1.

Examples of a PD: Bernoulli RVs

Bernoulli distribution: The outcome is either a "failure" (0) or a "success" (1).

 $\blacksquare X$ is a bernoulli RV when

$$Pr(X = 0) = p$$

 $Pr(X = 1) = (1 - p)$

For example, the USD-INR rises or not at the end of the day.

The bernoulli distribution has one parameter, *p*.

The CD of a Bernoulli RV

We need to know what p of the RV is. Say p = 0.35.
The CD of this Bernoulli RV is:

Value	Probability
X≤0	0.35
$X \leq 1$	1.00

Binomial distribution: The outcome is a sum (s) of a set of bernoulli outcomes (n).

For example, the number of times USD-INR rose in the last 10 days?

$$\Pr(X = s) = \frac{n!}{(n-s)!s!} p^s (1-p)^{(n-s)}$$

Here, p is the probability that the USD-INR rose in a day. The assumption is that p is constant over these 10 days.

The binomial distribution has two parameters: p, n.

Explaining n!

n! is mathematical short-hand for the product 1 * 2 * 3 * 4 * ... (n - 2) * (n - 1) * n
Example: 5! = 1 * 2 * 3 * 4 * 5
Note: 0! = 1! = 1

Testing concepts: Binomial RV CD

A bernoulli event has a probability of 0.4. What is the CD of a binomial RV that is the number of times that success can be acheived in 4 trials?

Testing concepts: Binomial RV CD

A bernoulli event has a probability of 0.4. What is the CD of a binomial RV that is the number of times that success can be acheived in 4 trials?

Value	PD	CD	
0	0.1296	0.1296	
1	0.3456	0.4752	
2	0.3456	0.8208	
3	0.1536	0.9744	
4	0.0256	1.0000	

Testing concepts: A discrete RV

A discrete RV has the following PD:

Х	1	2	4	5	8
Pr(X)	0.20	0.25		0.30	0.10

- 1. Find Pr(4).
- 2. Find ((Pr(x) = 2) or (Pr(x) = 4)).
- 3. Find $Pr(x \le 4)$.
- 4. Find Pr(x < 4).

Testing concepts: A discrete RV

A discrete RV has the following PD:

Х	1	2	4	5	8
Pr(X)	0.20	0.25		0.30	0.10

- 1. Find Pr(4). 0.15
- 2. Find ((Pr(x) = 2) or (Pr(x) = 4)). 0.55
- 3. Find $Pr(x \le 4)$. 0.60
- 4. Find Pr(x < 4). 0.45

Probability density functions

What is a probability density function?

- The probability density function (PDF) is the PD of a continuous random variable.
- Since continuous random variables are uncountable, it is difficult to write down the probabilities of all possible events.
 - Therefore, the PDF is always a function which gives the probability of one event, x.
- If we denote the PDF as function f, then

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

A probability distribution will contain all the outcomes and their related probabilities, and the probabilities will sum to 1.

The problem with estimating PDFs

- In a set of continuous random variables, the probability of picking out a value of exactly x is zero.
- We define the Pr(X = x) as the following difference:

$$\Pr(X \le (x + \Delta)) - \Pr(X \le x)$$

as Δ becomes an infinitesimally small number.

Here, $Pr(X \le x)$ is the cumulative density function of X.

tion (CDF)?

- Analogous to the discrete RV case, the CDF is the cumulation of the probability of all the outcomes upto a given value.
- Or, the CDF is the probability that the RV can take any value less than or equal to X.
- If we assume that the RV X can take values from $-\infty$ to ∞ , then theoretically,

$$F(X) = \int_{-\infty}^{X} f(x)d(x)$$

Reformulating the PDF in calculus

• Pr(x = X) is given as:

$$Pr(x = X) = F(X + \Delta) - F(X)$$
$$= d(F(x))/d(x)$$

for infinitesimally small Δ .

For continuous RVs, we approach the Pr(x) as the derivative of the CDF.

Uniform distribution: The outcome is any number that can take a value between a minimum (A) and a maximum (B) with equal probability.

For a uniform RV,

$$\Pr(X = x) = 1/(B - A)$$

The uniform density has two parameters, A, B.

Testing concepts: Uniform density CD

What is the form of the Uniform CDF, given that the maximum = B and minimum = A?

Testing concepts: Uniform density CD

What is the form of the Uniform CDF, given that the maximum = B and minimum = A?

$$F(X) = \int_{\text{minimum}}^{X} f(x)d(x)$$

=
$$\int_{A}^{X} \frac{1}{(B-A)d(x)}$$

=
$$\frac{(X-A)}{(B-A)}$$

ExamplesofaPDF:Nor-mal/Gaussian RV

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}((x-\mu)/\sigma)^2}$$

The normal has two parameters, μ , σ .

A lot of economic variables are assumed to be normal distributed.

Features of the normal PDF

- **RV**s can take values from $-\infty$ to ∞ .
- It is symmetric: Pr(-x) = Pr(x)
- When X is a normal RV with parameters μ , σ , then Y = 5.6 + 0.2X will also be a normal RV, with known parameters $(5.6 + 0.2\mu), (0.2\sigma)$.
- Note: Special case of a normal distribution is $\mu = 0, \sigma = 1$. This is called a standard normal distribution.

Probability distributions



Problems to be solved

Q1: Binomial RVs

A door-to-door salesperson has found that her success rate in selling is 0.2. If the salesperson contacts three persons, what is the probability that

- 1. She sells to all three?
- 2. She sells to exactly one?
- 3. She sells to at least one?

Q2: Uniform density CD

- If B = 10 and A = 5
 - 1. What is Pr(x = 5.5)?
 - 2. What is Pr(x = 4.5)?
 - 3. What is Pr(x > 7)?





References

 Chapter 2, SHELDON ROSS. Introduction to Probability Models. Harcourt India Pvt. Ltd., 2001, 7th edition