MA-120 Probability and Statistics

Nazar Khan
PUCIT

Lecture 10: Probability Distributions – Discrete

Random Variables

- A variable that takes on random values is called a **random variable**.
  - Outcome of a die is a random variable.
  - The next day of the week is <u>not</u> a random variable.
- A random variable assigns a value to the outcome of a random experiment.
- Random variables come in 2 types
  - Discrete – set of outputs is real valued, <u>countable</u> set
  - Continuous – set of outputs is real valued, <u>uncountable</u> set
- In this lecture: <u>discrete random variables and their probability distributions</u>.

Assigning Values to Outcomes

- Consider the sample space, S, consisting of all possible outcomes of four tosses of a coin.

$$S = \left\{ \begin{array}{cccc} HHHH & HHHT & HHTH & HTHH \\ THHH & HHTT & HTHT & THHT \\ TTHH & THTH & HTTH & TTTH \\ TTHT & THTT & HTTT & TTTT \end{array} \right\}$$

- Which outcome of this sample space do you like the most?
- What if I give you a 1 free shirt for every Head?

Random Variable

- What if I give you a 1 free shirt for every Head?

$$\begin{array}{cccc} HHHH \to 4 & HHHT \to 3 & HHTH \to 3 & HTHH \to 3 \\ THHH \to 3 & HHTT \to 2 & HTHT \to 2 & THHT \to 2 \\ TTHH \to 2 & THTH \to 2 & HTTH \to 2 & TTTH \to 1 \\ TTHT \to 1 & THTT \to 1 & HTTT \to 1 & TTTT \to 0. \end{array}$$

- Now you have ranked every outcome ω.
- When you assign a number to each outcome ω, you get a random variable.

Random Variable

Let X=number of free shirts that you win
Same as X = number of heads in 4 coin tosses

$$
\begin{array}{llll}
HHHH \to 4 & HHHT \to 3 & HHTH \to 3 & HTHH \to 3 \\
THHH \to 3 & HHTT \to 2 & HTHT \to 2 & THHT \to 2 \\
TTHH \to 2 & THTH \to 2 & HTTH \to 2 & TTTH \to 1 \\
TTHT \to 1 & THTT \to 1 & HTTT \to 1 & TTTT \to 0.
\end{array}
$$

$$
\{X = 2\} = \left\{ \begin{array}{lll} HHTT, & HTHT, & THHT \\ TTHH, & THTH, & HTTH \end{array} \right\}
$$

P(X=2) = 6/16
X is a <u>variable</u> that takes on <u>random</u> values between 0 and 4.

Random Variable

Experiment: Roll a fair die twice and note the 2 faces that come up. All possible outcomes are

$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(1,5)$	$(1,6)$
$(2,1)$	$(2,2)$	$(2,3)$	$(2,4)$	$(2,5)$	$(2,6)$
$(3,1)$	$(3,2)$	$(3,3)$	$(3,4)$	$(3,5)$	$(3,6)$
$(4,1)$	$(4,2)$	$(4,3)$	$(4,4)$	$(4,5)$	$(4,6)$
$(5,1)$	$(5,2)$	$(5,3)$	$(5,4)$	$(5,5)$	$(5,6)$
$(6,1)$	$(6,2)$	$(6,3)$	$(6,4)$	$(6,5)$	$(6,6)$

which correspond to a simple sample space.

Let X=sum of the two die rolls. All possible outcomes are now S_1={2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12} which is not a simple sample space.

Values of X	2	3	4	5	6	7	8	9	10	11	12
Probabilities	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Random Variable

Values of X	2	3	4	5	6	7	8	9	10	11	12
Probabilities	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Find P(X>=9)?

Find P(4≤X<8)?

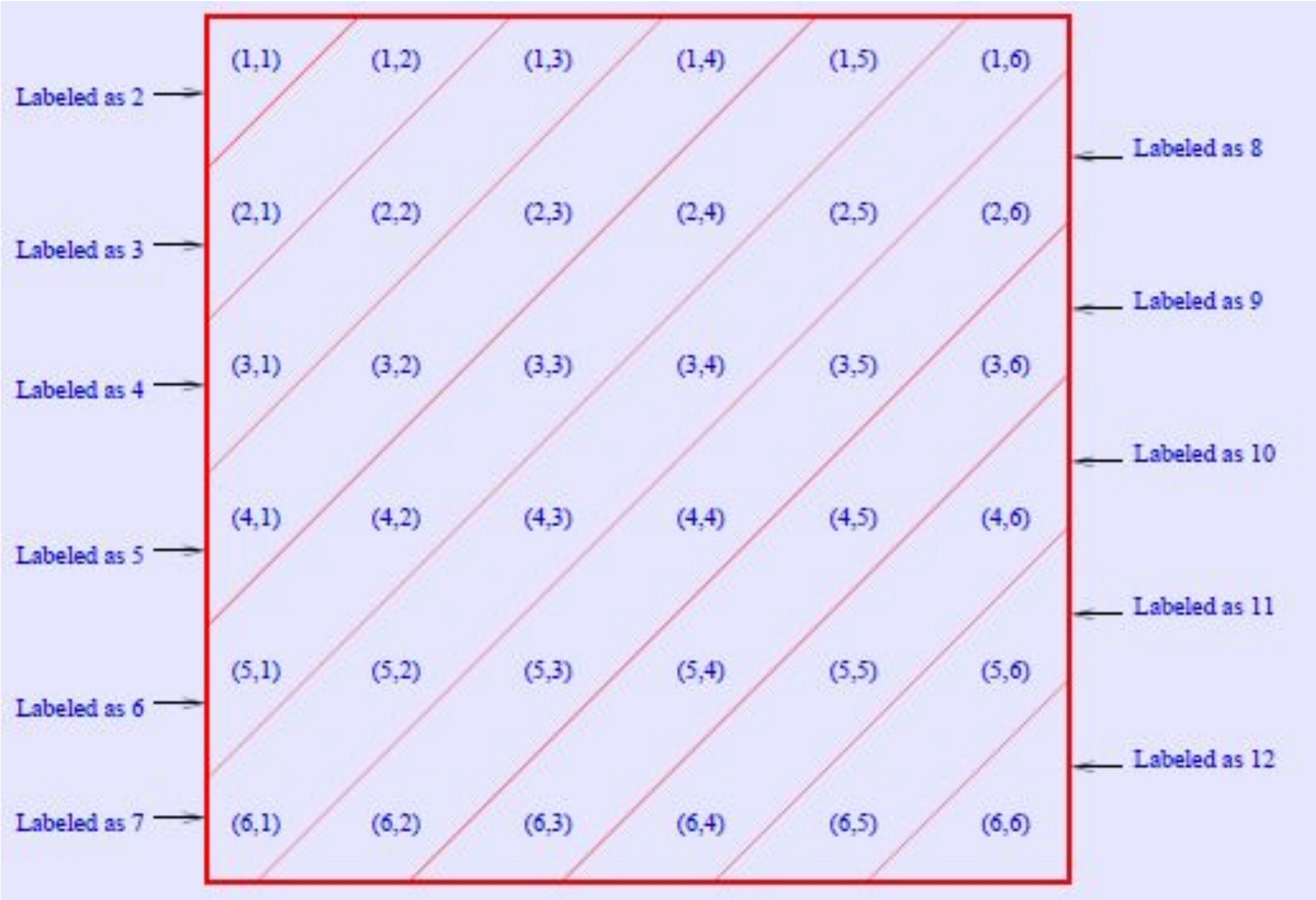
Random Variable

- A random variable does 2 things:
  1) partitions the sample space into disjoint events
  2) labels the elements of these events by appropriate numbers corresponding to their rankings

- If the labels are discrete, then the random variable is discrete.

- If the labels are continuous, then the random variable is continuous.

$$\{X=2\}=\{(1,1)\}$$
$$\{X=3\}=\{(1,2),(2,1)\}$$
$$\{X=4\}=\{(1,3),(2,2),(3,1)\}$$
$$\{X=5\}=\{(1,4),(2,3),(3,2),(4,1)\}$$
$$\{X=6\}=\{(1,5),(2,4),(3,3),(4,2),(5,1)\}$$
$$\{X=7\}=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$
$$\{X=8\}=\{(2,6),(3,5),(4,4),(5,3),(6,2)\}$$
$$\{X=9\}=\{(3,6),(4,5),(5,4),(6,3)\}$$
$$\{X=10\}=\{(4,6),(5,5),(6,4)\}$$
$$\{X=11\}=\{(5,6),(6,5)\}$$
$$\{X=12\}=\{(6,6)\}$$



Partition of S induced by the random variable X

Binomial Experiments

Example

We toss a fair coin twice. Let A = {H occurs on the first toss} and let B = {H occurs on the second toss}. Are A and B independent?

The sample space is S = {HH, HT, TH, TT}. It is simple since the coin is fair.
Now A = {HH,HT}, B = {HH, TH}, $A \cap B$ = {HH}.
Therefore, $P(A) = 1/2 = P(B)$ and $P(A \cap B) = 1/4$.
A and B are independent events since $P(A \cap B) = 1/4 = 1/2 \cdot 1/2 = P(A)P(B)$.

Binomial Experiments

Example

Suppose now <u>we toss an unfair coin twice</u>, so that the probability of seeing H on a toss is *p* and it is not necessarily equal to 1/2 . How can we compute probabilities for the singleton sets of the sample space S = {HH, HT, TH, TT}?

P({HH})=p*p   P({HT})=p*(1-p)

P({TT})=(1-p)*(1-p) P({TH})=(1-p)*p

Binomial Experiments

Example

Suppose now we toss an unfair coin thrice. How can we compute probabilities for the singleton sets of the sample space S?

S=?   P(ω ∈ S)=?

Binomial Experiments

Example

Now let's count the number of heads in the 3 tosses.

P(0 heads in 3 tosses) = $p(\{TTT\}) = (1-p)^3$

P(1 heads in 3 tosses)

$= p(\{HTT \cup THT \cup TTH\})$

$= p(HTT) + p(THT) + p(TTH)$

$= p(1-p)^2 + p(1-p)^2 + p(1-p)^2 = 3p(1-p)^2 = C(3,1)p(1-p)^2$

P(2 heads in 3 tosses)

$= p(\{HHT \cup HTH \cup THH\})$

$= p(HHT) + p(HTH) + p(THH)$

$= p^2(1-p) + p^2(1-p) + p^2(1-p) = 3p(1-p)^2 = C(3,2)p(1-p)^2$

P(3 heads in 3 tosses) = $p(\{HHH\}) = p^3$

P(j heads in 3 tosses) = $C(3,j)p^j(1-p)^{3-j}$

P(j heads in n tosses) = $C(n,j)p^j(1-p)^{n-j}$

Terminology

Bernoulli experiment: An experiment which has only two possible outcomes is called a Bernoulli experiment. One of the outcomes is called Heads or "Success" and the other Tails or "Failure".
Bernoulli process: An experiment consisting of *n* unrelated (independent) and identical Bernoulli experiments.

Product Spaces & Independent Trials

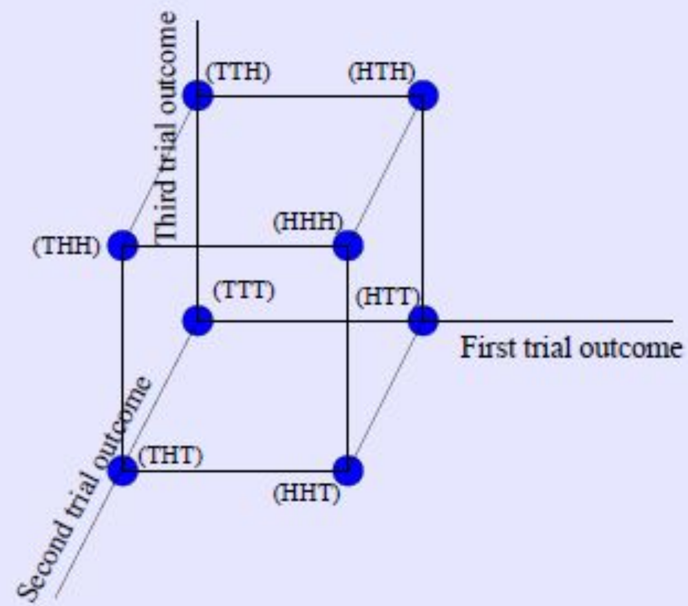


FIGURE 2.1: Sample Space as Set Product $S = S_1 \times S_2 \times S_3$.

- If a random experiment consists of *m* sub-experiments, each called a trial, the sample space, S, may be considered a <u>set product</u> of the individual sets, S_i, i = 1, 2, · · · ,m, consisting of the outcomes of each trial.
- $S = S_1 \times S_2 \times S_3 \times \cdots \times S_m$.

Product Spaces & Independent Trials

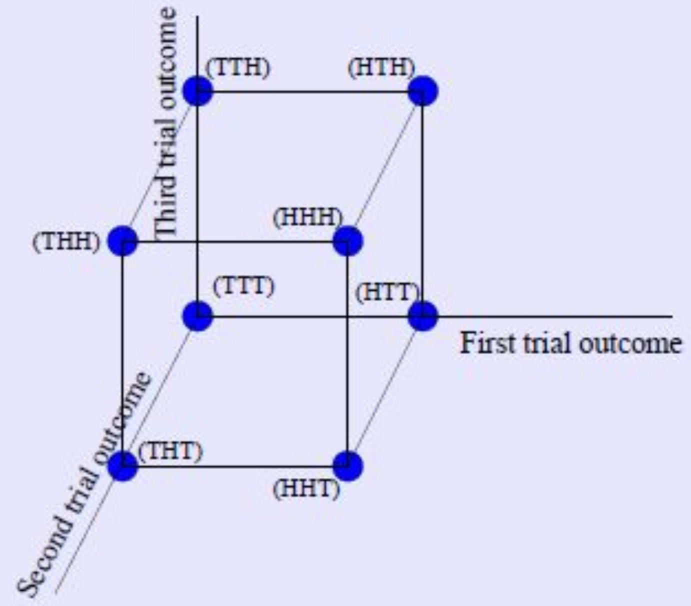


FIGURE 2.1: Sample Space as Set Product $S = S_1 \times S_2 \times S_3$.

- If a random experiment consists of *m* sub-experiments, each called a trial, the sample space, S, may be considered a <u>set product</u> of the individual sets, S_i, $i = 1, 2, \cdots, m$, consisting of the outcomes of each trial.
- $S = S_1 \times S_2 \times S_3 \times \cdots \times S_m$.

“A sequence of Bernoulli trials”

“A sequence of Bernoulli trials” means that

- the trials are independent,
- each trial can have only 2 possible outcomes, (called Heads and Tails, denoted by H and T), and
- the probability of an H remains the same from trial to trial, and is usually denoted by p.

Binomial Experiments

Example

We toss a fair coin twice. Let A = {H occurs on the first toss} and let B = {H occurs on the second toss}. Are A and B independent?

The sample space is S = {HH, HT, TH, TT}. It is simple since the coin is fair.
Now A = {HH,HT}, B = {HH, TH}, $A \cap B$ = {HH}.
Therefore, $P(A) = 1/2 = P(B)$ and $P(A \cap B) = 1/4$.
A and B are independent events since $P(A \cap B) = 1/4 = 1/2 \cdot 1/2 = P(A)P(B)$.

Geometric experiment

2. Geometric experiment. Over the entire (infinite) sequence of Bernoulli trials, what is the probability that k tails will be observed to get the first head?

$p(1 - p)^k$, for $k = 0, 1, 2, \cdots$.

Note the difference between the Binomial model and the Geometric model.

Shifted geometric experiment

3. Shifted geometric experiment. Over the entire (infinite) sequence of Bernoulli trials, what is the probability that k trials will be needed to get the first head?

$p(1-p)^{k-1}, \quad \text{for } k = 1, 2, 3, \cdots.$

The justification is quite similar to the one used in the geometric experiment. The difference being that now we count the trial that results into the first head.

Negative binomial experiment

4. Negative binomial experiment. Over the entire (infinite) sequence of Bernoulli trials, what is the probability that k tails will be observed to get the first 10 heads?

$$\binom{k+10-1}{10-1} p^{10}(1-p)^k \qquad \text{for } k = 0, 1, 2, \cdots$$

H.W: See if you can justify this result.

Shifted negative binomial experiment

5. Shifted negative binomial experiment. Over the entire (infinite) sequence of Bernoulli trials, what is the probability that k trials will be observed to get the first 10 heads?

$$\binom{k-1}{10-1} p^{10}(1-p)^{k-10} \qquad \text{for } k = 10, 11, 12, \cdots$$

H.W: See if you can justify this result.

Poisson

Multinomial

Hypergeometric

Example

Three dice are rolled.
You select a number from the set {1, 2, 3, 4, 5, 6}, (say 4).
If no 4 comes on the three dice you loose 100 Rupees.
Otherwise you win 100*(number of 4s)
Suppose you decide to play this game once. What are your chances of (i) loosing 100 Rupees, (ii) winning 100 Rupees, (iii) winning 200 Rupees, (iv) winning 300 Rupees?
Let X be the random variable representing your net gain (or loss) after one game. What is the density of X?

Example

Three dice are rolled.
You select a number from the set {1, 2, 3, 4, 5, 6}, (say 4).
If no 4 comes on the three dice you loose 100 Rupees.
Otherwise you win 100*(number of 4s)
Suppose you decide to play this game once. What are your chances of (i) loosing 100 Rupees, (ii) winning 100 Rupees, (iii) winning 200 Rupees, (iv) winning 300 Rupees?
Let X be the random variable representing your net gain (or loss) after one game. What is the density of X?

Summary

- Probability
  - Random Experiment, Outcome, Sample space, Event, P(event)$\in$[0,1]
- Independence
  - Joint=product of marginals (P(AB)=P(A)P(B))
- Random Variables
  - Assign numbers to outcomes
  - Can be discrete (countable set) or continuous
- Discrete Random Variables
  - Binomial
  - Geometric and shifted
  - Negative Binomial and shifted
  - Poisson
  - Hyper-geometric
  - Multinomial
  - …