TRAINING COURSES ON APPLIED ECONOMETRIC ANALYSIS (SUMMER SCHOOL) FOR YOUNG ECONOMISTS / RESEARCHERS 2016-2017

Basics of Probability theory

Farrukh Ataev

14 September 2016

Outline

- **Session 3 Sets and Basic probability concepts**
  - (Wed 14 Sept at 13:30-15:30)
- **Session 4 Probability distributions (Discrete)**
  - (Wed 14 Sept at 15:30-17:00)
- **Session 1 Probability distributions (Continuous)**
  - (Thu 15 Sept at 9:00-10:30)
- **Session 2 Inferential statistics and Confidence intervals**
  - (Thu 15 Sept at 11:00-12:30)

Session 4 Discrete probability distribution

- Random variable (discrete and continuous)
- Probability density function (PDF)
- Probability distribution (discrete and continuous)
- Discrete probability distribution
  - Binomial probability distribution
  - Poisson probability distribution

Random variable

- A numerical value related to an outcome of an experiment
- Can be discrete or continuous
- Answers the question “how many?” or “how much?”

Examples:

1) Toss a coin twice. The random variable X is a number of heads. Then X = {0,1,2}
2) A student is taking an exam. The random variable X defines the student’s mark on a scale of 0 to 100. Then $0 \le X \le 100$.

Exercise:

1) A die is rolled 3 times. X defines how many times an even number occurs. Find the range of X.

2) The customer expenditures at a grocery store are observed. The random variable X defines a customer’s expenditure. Find the range of X.

Probability function and probability distributions

The **probability function**, denoted by $f(x)$, provides the probability of occurrence of a random variable. The required conditions for discrete probability function are $f(x) \geq 0$ and $\sum f(x) = 1$.

A **probability distribution** is a table, graph, or mathematical formula that shows all possible values of the random variable x and the associated probability function $f(x)$.

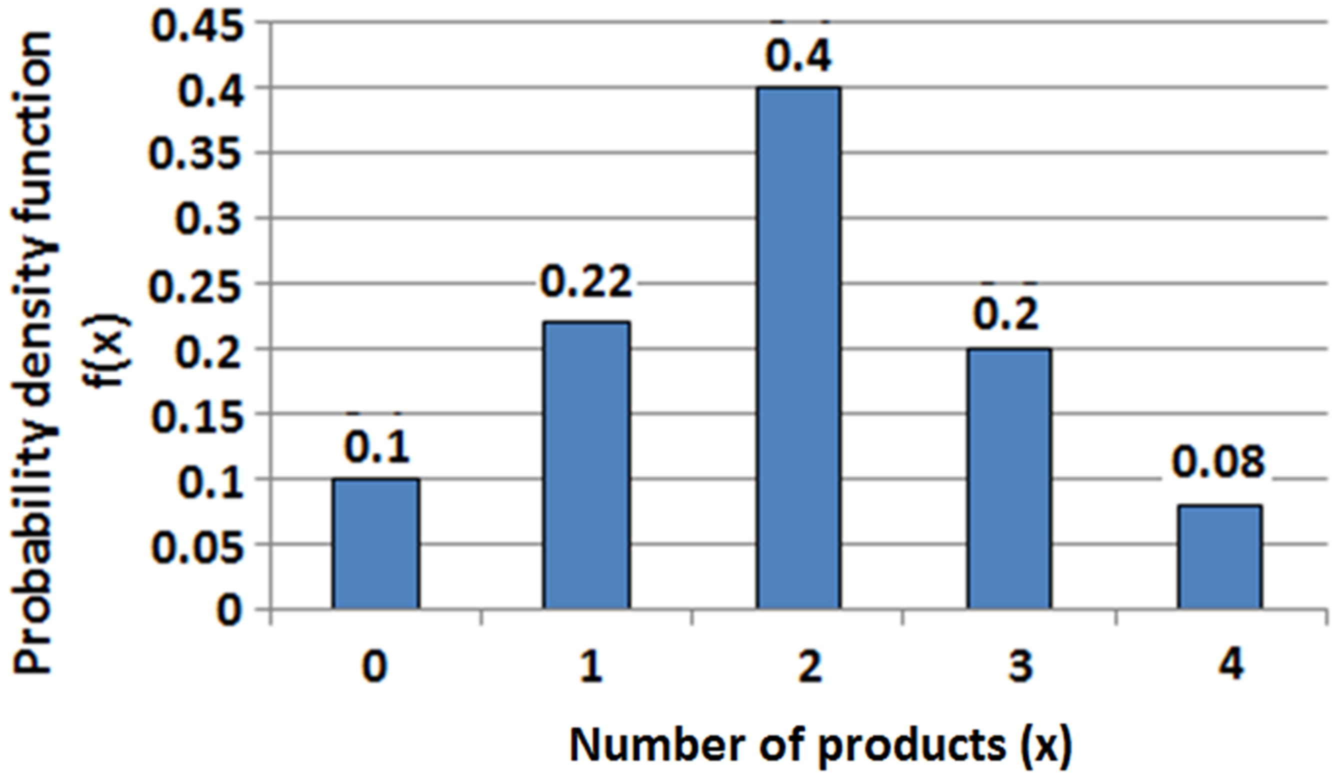
Two types of a probability distribution:

1) Discrete probability distribution (Binomial, Poisson, hypergeometric);
2) Continuous probability distribution (Uniform, Normal, Exponential).

Example 1 (Discrete probability distribution)

Number of products purchased by customers:

# of products x	# of customers f	P (X=x)	f (x)
0	10	0.10	0.10
1	22	0.22	0.22
2	40	0.40	0.40
3	20	0.20	0.20
4	8	0.08	0.08
	100	**1**	**1**

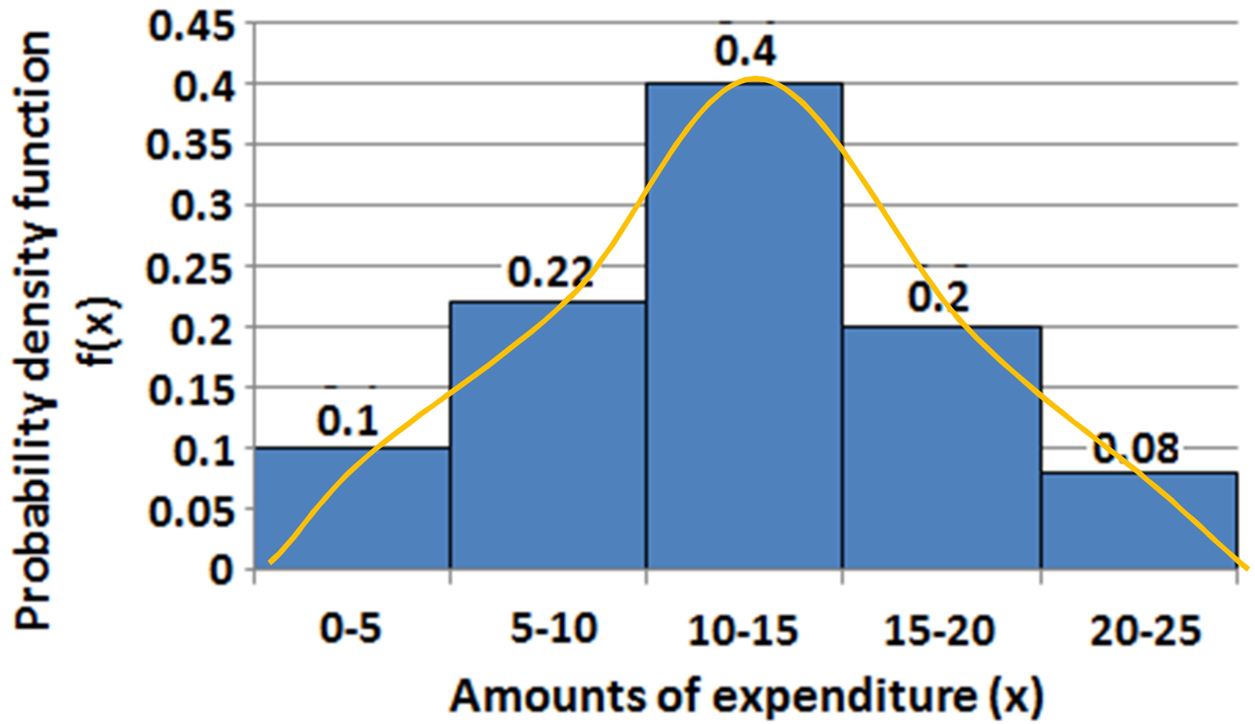


Exercise: Find **1)** $P(X=2) = f(2)$, **2)** $P(1 \leq X \leq 3)$, **3)** $P(X<4)$

Example (Continuous probability distribution)

Amounts of expenditure by customers:

x	# of cust.	P (X)
0-5	10	0.10
5-10	22	0.22
10-15	40	0.40
15-20	20	0.20
20-25	8	0.08
	100	**1**



Exercise: Find **1)** P(15≤X<20), **2)** P(X=7), **3)** P(8≤X<12)

Expected value (mathematical expectation)

Number of products purchased by customers:

# of products x	# of customers f	P (X=x)	f (x)	fx	xp
0	10	0.10	0.10	0	0
1	22	0.22	0.22	22	0.22
2	40	0.40	0.40	80	0.80
3	20	0.20	0.20	60	0.60
4	8	0.08	0.08	32	0.32
	100	**1**	**1**	**194**	**1.94**

Mean = $\bar{x} = \frac{\sum fx}{\sum f} = \frac{194}{100} = 1.94$

E(X) = $\sum xp = \sum xf(x) = 1.94$

Variance = $s^2 =$

$$\frac{\sum f(x-\bar{x})^2}{\sum f} = \frac{113.64}{100} = 1.1364$$

V(X) = $\sum\left(X - E(X)\right)^2 P(X) =$
1.1364

Example of Mathematical expectation

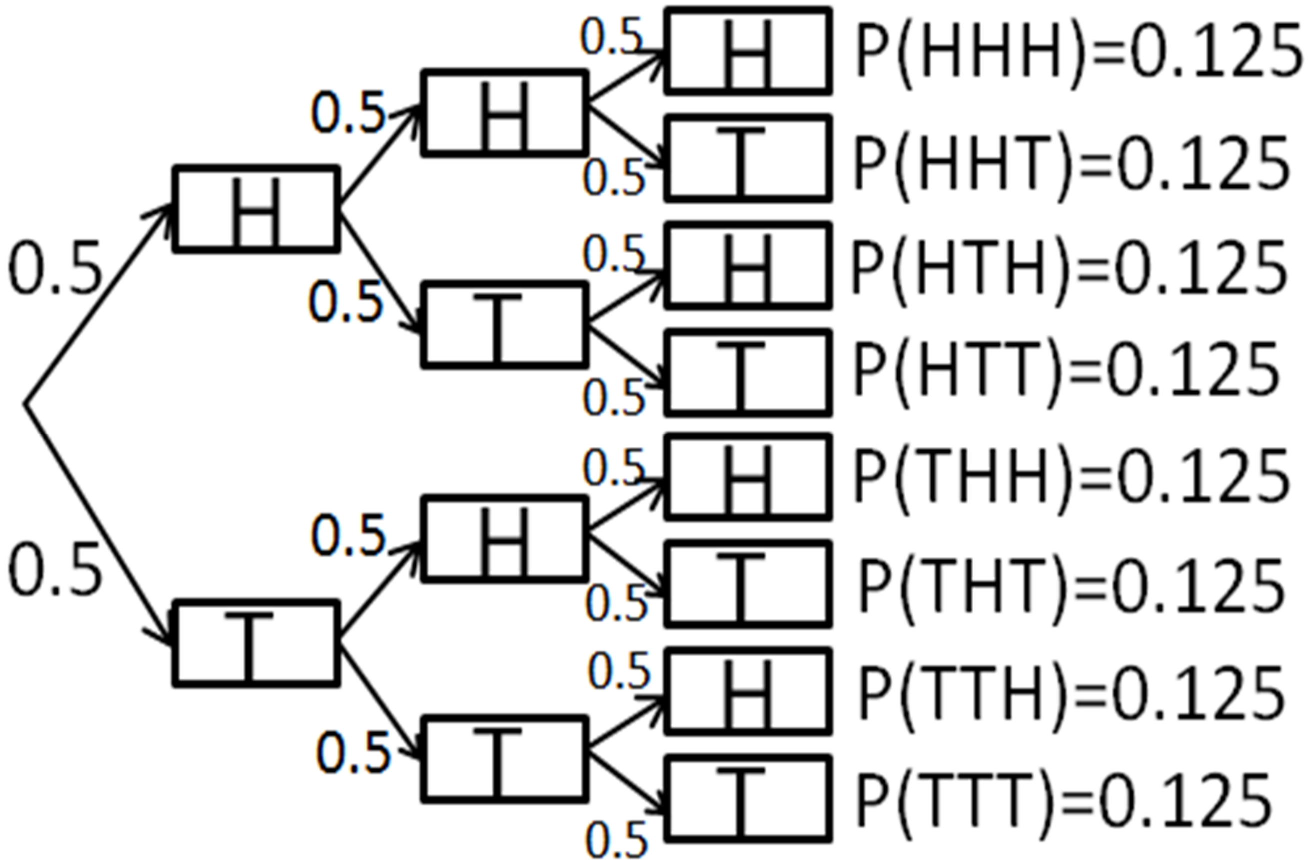
In the game of rolling a die, a player wins $10 if he gets 1, $20 if 2, $30 if 3, $40 if 4, $50 if 5, but he loses $120 if he gets 6. Find the expected sum of the money won (or mean value).

X	$10	$20	$30	$40	$50	– $120
P	1/6	1/6	1/6	1/6	1/6	1/6

$$E(X) = \sum XP(X) = 10 \cdot \frac{1}{6} + 20 \cdot \frac{1}{6} + 30 \cdot \frac{1}{6} + 40 \cdot \frac{1}{6} + 50 \cdot \frac{1}{6} - 120 \cdot \frac{1}{6} = 5$$

Example 2 (Discrete probability distribution)

A **discrete probability distribution** of a coin tossed 3 times and x times appearing “Head”



x	**f (x)**	**F(x)**
0	0.125	0.125
1	0.375	0.5
2	0.375	0.875
3	0.125	1
	1	

Exercise: Find **1)** $P(X=2) = f(2)$, **2)** $P(1 \le X \le 3)$, **3)** $P(X<3)$

Binomial experiment and Binomial distribution

❖ Shows the number of occurrences in a multiple step experiment

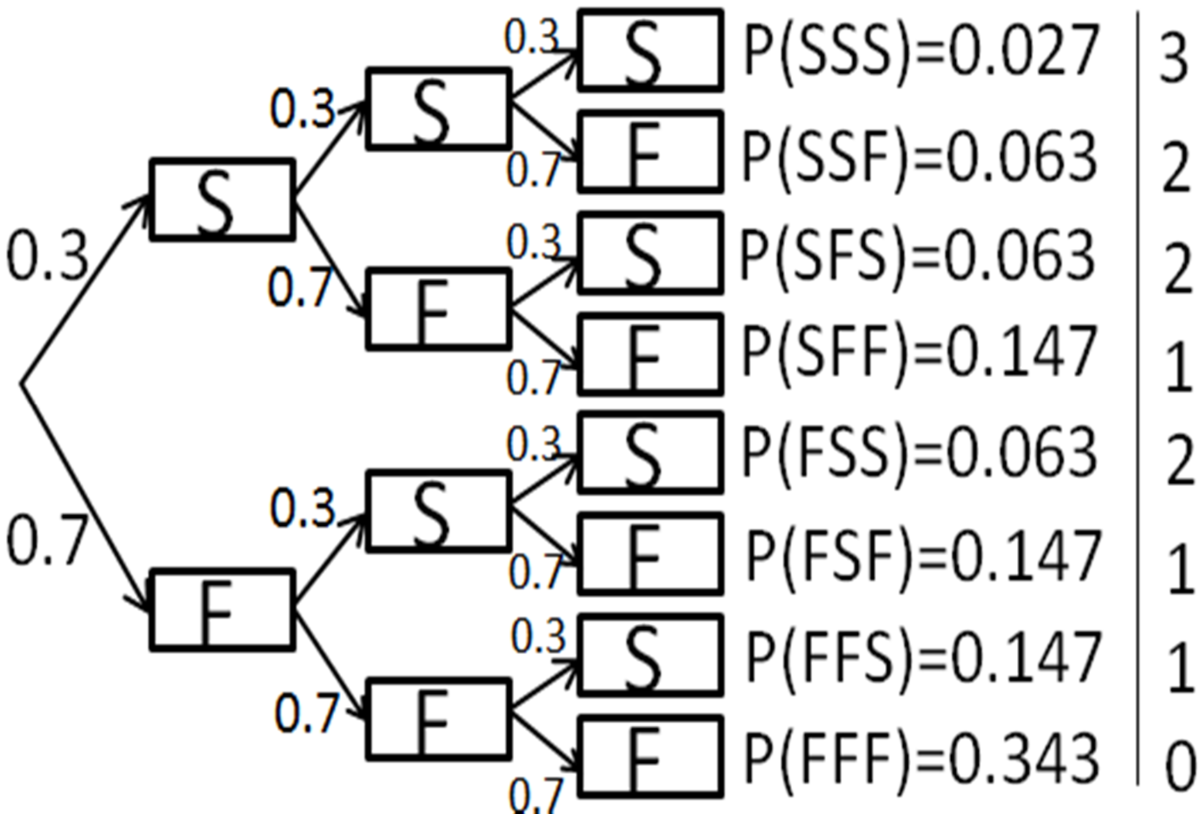
Binomial experiment has 4 assumptions:

1) It has a sequence of n trials
2) Two outcomes are possible:
   Success (p) and Failure (1 – p)
3) The probability of a success does not change from trial to trial.
4) Trials are independent

Example: Tossing a coin 3 times and counting heads is a binomial experiment.

Binomial experiment (distribution) I

During an hour 3 customers enter a store and the probability of a customer purchase is 0.3. What is the probability of 2 out of 3 customers purchasing?



x	**f (x)**	**F(x)**
0	0.343	0.343
1	0.441	0.784
2	0.189	0.973
3	0.027	1
	1	

Binomial experiment (distribution) II

1-method: $f(x) = C_x^n p^x (1-p)^{n-x}$

$$f(2) = C_2^3 p^2 (1-p)^{3-2}$$

$$= \frac{3!}{2!\,(3-2)!} 0.3^2 (1-0.3) = 0.189$$

2-method:

Binomial Table: n = 3, x = 2, p = 0.3

f (2) = P (X = 2) = 0.189

x	**f (x)**	**F(x)**
0	0.343	0.343
1	0.441	0.784
2	0.189	0.973
3	0.027	1
	1	

Cum. Bin. Table: n = 3, x = 2, x = 3, p = 0.3

f (2) = P(X ≤ 2) – P (X ≤ 1) = 0.973 – 0.784 = 0.189

3-method: In MS Excel: BINOM.DIST(x,n,p,0)

BINOM.DIST(2,3,0.3,0) = 0.189 or

BINOM.DIST(2,3,0.3,1) – BINOM.DIST(1,3,0.3,1) = 0.189

Poisson experiment and Binomial distribution

- Shows the number of occurrences over a certain time or space

Poisson experiment has 2 assumptions:

1) The probability of an occurrence is the same for any two intervals of equal length
2) Occurrences in the intervals are independent

Examples:

1) The number of arrivals at a car wash in 1 hour
2) The number of repairs needed in 10 kilometers of highway
3) The number of leaks in 100 kilometers of pipeline.

Poisson experiment (distribution) I

Phone calls arrive at the rate of 4 calls per 5 minutes at the reservation desk for UzAir company. Find the probability of receiving 3 calls in a 5 minute interval.

This is a Poisson experiment, because:

1) The probability of phone calls (4) is the same for any two periods of equal length
2) Occurrence or non-occurrence of a phone call in any period is independent of the occurrence or non-occurrence in any other period.

Poisson experiment (distribution) II

1-method: $f(x) = \dfrac{\mu^x e^{-\mu}}{x!}$

$$f(3) = \frac{4^3 e^{-4}}{3!} = \frac{64 \cdot 0.0183}{6} = 0.1954$$

2-method: Poisson Table: $\mu = 4$, x = 3

f (3) = P (X = 3) = 0.1954

Cum. Poisson Table:

f (3) = P(X ≤ 3) – P (X ≤ 2) = 0.4335 – 0.2381 = 0.1954

3-method: In MS Excel: POISSON.DIST(x,μ,0)

POISSON.DIST(3,4,0) = 0.1954

POISSON.DIST(3,4,1) – POISSON.DIST(2,4,1) = 0.1954

Thank you