Kyiv School of Economics
April 24, 2010

Preparatory Course in Mathematics

Lecture 3

1 Probability Theory

1.1 Major definitions

We define several types of observable events :

- **Sure event:** events that will *definitely* happen (under a certain set of conditions);
- **Impossible event:** event that will *definitely not* happen;
- **Random event:** may or may not happen: the *outcome* of such events cannot be predicted with certainty.

Example. *When you heat water, it will* ***surely*** *boil at the temperature of* 100 °C*. However, when you are choosing between buying 2 goods, the outcome (first, second or both items are purchased) is not so certain.*

Definition. Under **random events** we understand outcomes of a random experiment so that:

a) outcomes are multiple;

b) outcomes are uncertain;

c) experiment can be repeated identically.

Example.

- *tossing coins (tails or heads?)*
- *measuring annual national GDP (we are unsure about the exact amount beforehand)*
- *hearing "will you marry me?" from your boyfriend*
- *hearing "yes" if you propose to your girlfriend*

We assume that **elementary events** are:

- **incompatible:** only one outcome is possible (either heads or tails can be shown);
- **equally possible:** (numbers on the equally balanced dice, sides of a coin, etc.)

Example. *Two dice are thrown. What is a probability that the sum is 3?*

Solution (wrong). Possible logic: **2** outcomes are possible: (a) $S = 3$, and (b) S$\neq$ 3. Therefore $p = \frac{1}{2}$. This logic is wrong, since outcomes (a) and (b) are not equally possible.
Solution (correct). Total number of equally possible outcomes is **36**: $\{1, 1\}, \{1, 2\}, \{1, 3\}, \ldots, \{6, 6\}$. Only two of them lead to event $S = 3$: $\{1, 2\}$ and $\{2, 1\}$. Therefore $p = \frac{2}{36} = \frac{1}{18}$.

Definition (Classical)**. Probability** of an event is defined as the relative frequency: ratio of the favorable outcomes m to the total number of all possible outcomes n:

$$P = \frac{m}{n}$$

Example. *Event H is "After a coin toss, heads are shown". What is $P(H)$?*

Solution. Intuitively, the solution is $\frac{1}{2}$. However, the precision of the result depends on the number of experiments. Let's assume we have thrown the same coin 10 times. Here are the outcomes:

T H T H H H T T H H

If we base our solution on the results of this experiment only, then we have to conjecture that $P(H) = \dfrac{6}{10}$. The more often an experiment is repeated, the more precise the probability is:

$$\text{If } n \to \infty, P(H) = P(T) \to \frac{1}{2}$$

.

Definition. Now we define Ω (sample space) as the set of **all** possible outcomes. When we talk about some event, we presume this event belongs to the sample space: $P(A)$: probability of event $A \subseteq \Omega$.

Example. *We toss a coin 3 times. Describe Ω, probability of any outcome, probability of two heads.*

Solution.

1. Ω ={HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

2. Probability of every outcome: $P(A) = \dfrac{1}{8}$

3. Probability of two heads: B=Two heads. B= $\{HHT, HTH, THH\}$ P(B)=$\dfrac{3}{8}$

Definition. Given that, we define the following:

1. **sure event:** P(A)=1

2. **impossible event:** P(A)=0

3. **uncertain event:** $0 < P(A) < 1$

Example. *In a group of 10 students, 7 are females. What is a probability that among 6 randomly picked students there are 4 females?*

Solution.

$$P = \frac{m}{n}$$

$$n = C_{10}^{6} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 210$$

$$m = C_7^4 \cdot C_3^2 = \frac{7 \cdot 6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3 \cdot 4} \cdot \frac{3 \cdot 2}{1 \cdot 2} = 105$$

(4 females out of 7 in the sample, 6-4=2 males out of 10-7=3 males)

$$P = \frac{m}{n} = \frac{105}{210} = \frac{1}{2}$$

1.2 Properties of Probabilities

Sum of probabilities

$P(A + B) = P(A) + P(B)$ (given that A and B incompatible)

Example. *There are 30 balls: 10 red, 5 blue, 15 white. C =colored ball. $P(C)$-?*

Solution. $P(Red) = \dfrac{10}{30}$, $P(Blue) = \dfrac{5}{30}$. Therefore $P(C) = \dfrac{10+5}{30} = \dfrac{1}{2}$

Opposite events

Definition.

Even $\bar{A}$ is **opposite** to event A if $A + \bar{A} = \Omega$

Therefore, $P(A) + P(\bar{A}) = 1$

Example. *There are m females in a group of n students. If we randomly pick k students, what is a probability that at least one of them is a female?*

Solution.

$P(A) = 1 - P(\bar{A})$ $\bar{A}$: no females at all $P(\bar{A}) = \dfrac{C_{n-m}^k}{C_n^k}$ *Note*: There are $n - m$ males in the group.

$\Rightarrow P(A) = 1 - \dfrac{C_{n-m}^k}{C_n^k}$

Dependent and Independent Events

Example. *The coin is tossed twice. Event HA= "Heads are in the first throw, anything else in the second". AH="Head are in the second throw.". It turns out that $P(HA)$ is independent of $P(AH)$. Therefore these two events are considered **independent**.*

Definition. Events are **independent** of probability of one *does not* depend on occurrence of another.

Example. *In the group of 100 students, there are 60 males and 40 females. We randomly pick one student, and then another one. What is the probability that the second student is a male?*

Solution. Probability of the second student's gender depends on the outcome of the first draw. Hence, these two events are **dependent**.

1. First student was a male. Then probability of the second student being a male is $P(M|M) = \dfrac{59}{99}$.
2. First student was a female. Then probability of the second student being a male is $P(M|F) = \dfrac{60}{99}$.

Product of two events

(Simultaneous occurrence of 2 events)

If A and B are *independent*, then $P(AB) = P(A) \cdot P(B)$.

Example. $P(HH) = P(H) \cdot P(H) = \dfrac{1}{2} \cdot \dfrac{1}{2} = \dfrac{1}{4}$.

Example. *There are three groups, 20 students in each.*
In Group 1: 7 females, in Group 2: 10 females, in Group 3: 15 females.
We randomply pick one student from each group. What is the probability that all three students are males?

Solution.

$P(M_1) = \dfrac{13}{20}$, $P(M_2) = \dfrac{10}{20}$, $P(M_3) = \dfrac{5}{20}$

$P(M_1 \cdot M_2 \cdot M_3) = \dfrac{13 \cdot 10 \cdot 5}{20^3} = 0.08125$

If events A and B are *dependent*:
$P(AB) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

Example. *In a group of 100 students, there are 60 males and 40 females. We randomly pick two students, one after another.*

$P(M) = 0.6, P(F) = 0.4$

$P(MM) = P(M) \cdot P(M|M) = 0.6 \cdot \dfrac{59}{99} = \dfrac{60}{100} \cdot \dfrac{59}{99}$

$P(MF) = P(M) \cdot P(F|M) = 0.6 \cdot \dfrac{40}{99} = \dfrac{60}{100} \cdot \dfrac{40}{99}$

$P(FM) = P(F) \cdot P(M|F) = 0.4 \cdot \dfrac{60}{99} = \dfrac{40}{100} \cdot \dfrac{60}{99}$

$P(FF) = P(F) \cdot P(F|F) = 0.4 \cdot \dfrac{39}{99} = \dfrac{40}{100} \cdot \dfrac{39}{99}$

L="At least one female"

$P(L) = P(1f) + P(2f) = \dfrac{C_{40}^1 \cdot C_{60}^1}{C_{100}^2} + \dfrac{C_{40}^2}{C_{100}^2} = \dfrac{40 \cdot 60 \cdot 2}{100 \cdot 99} + \dfrac{40 \cdot 39}{100 \cdot 99}$

Probability sum of compatible events

Compatible events: one, another or both.
$P(A+B) = P(A) + P(B) - P(AB)$
If A and B are independent: $P(A) + P(B) - P(A) \cdot P(B)$
If A and B are dependent: $P(A) + P(B) - P(A) \cdot P(B|A)$

Example.
Probability of event A (passing the math exam) is $P(A) = 0.7$
Probability of event B (passing the economics exam) is $P(B) = 0.8$
$P(A+B) = 0.7 + 0.8 - 0.7 \cdot 0.8 = 0.94$ – *probability that at least one exam is passed*
$P(A) \cdot P(B|A) = P(A) \cdot P(B) = 0.7 \cdot 0.8 = 0.56$ – *probability that both exams are passed*

Bayes Formula

From the property $P(A \cap B) = P(AB) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

It follows that $P(A|B) = \dfrac{P(A) \cdot P(B|A)}{P(B)}$

In general,

$$\boxed{P(B) = P(C_1) \cdot P(B|C_1) + P(C_2) \cdot P(B|C_2)}$$

Example. *40% of population has some disease. All individuals can be administered a blood test. However, the results of the test can wrongly indicate the disease (false positive) in 20% cases, and wronly show the absence of the disease (false negative) in 10% cases. A random individual is administered the test, and the test is positive. What is a probability that the disease is present?*

Solution. A_1: an individual has the disease, $P(A_1)$=0.4

$A_2 = A_1^c$: an individual does not have the disease, $P(A_2) = 0.6$

B: the test indicates a disease.

$P(A_1|B) = \dfrac{P(A_1) \cdot P(B|A_1)}{P(B)}$

$P(A_1) = 0.4$

$P(B|A_1) = 1 - 0.1 = 0.9$ (disease is present and shown by the test)

$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) = 0.4 \cdot 0.9 + 0.6 \cdot 0.2$

$\Rightarrow P(A_1|B) = \dfrac{0.4 \cdot 0.9}{0.4 \cdot 0.9 + 0.6 \cdot 0.2} = 0.75$

1.3 Intro to Math Statistics

Two Types of Random Variables

1. **Discrete Random Variables**

   Example. *Let's return to the "Two dice example", but now write out the outcomes in the table. The first row shows all possible outcomes (the sum on the dice), the bottom row shows the probability of the event.*

x	2	3	4	5	6	7	8	9	10	11	12
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

   One can see that the total sum of probabilities is **always** one: $\sum_i P(x_i) = 1$

   In this example, x is *countable* and *finite*.

Example. *We keep tossing a coin. Let Y denote the number of tosses until the first head comes up.*

$P(Y = 1) = P(H) = \frac{1}{2}$

$P(Y = 2) = P(TH) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$P(Y = 3) = P(TTH) = \frac{1}{8}$

...

$P(Y = k) = \left(\frac{1}{2}\right)^k$

In this example, the variable Y is countable, but infinite.

If X is countable (or descrete) variable, we usually can find a probability that X is equal to a certain value, given the **probability mass function**:

$$P_i = P(X = x_i), \ i = 1, 2, \ldots$$

Example. *In our regular two dice example,*

$P(X = 5) = \frac{4}{36}$

If we are looking for *an interval of values* for X, we consider **cumulative distribution function**:

Example.

$$F(x) = P(X \leq 9) = \sum_{i=2}^{9} P(X = i) = \frac{30}{36} = \frac{5}{6}$$

Example. *A random variable X has the following distribution function:*

x	*1*	*2*	*4*
P(x)	*0.2*	*0.4*	*0.4*

The cumulative distribution function is:

$$F(x) = \begin{cases} 0, & \textit{if} \quad x < 1; \\ 0.2, & \textit{if} \quad 1 \leq x < 2; \\ 0.6, & \textit{if} \quad 2 \leq x < 4; \\ 1, & \textit{if} \quad x \geq 4; \end{cases}$$

2. **Continuous Random Variables**

It is possible that the random variable X is continuous, or takes so many values that we can only talk about some *subintervals*, rather than point estimations. Therefore, for continuous random variables, $P(X = x) \to 0$.

Example. *Time is a clear example of a continuous variable. Let's consider a person waiting for a bus. The probability of her waiting for exactly 5 minutes is essentially zero:*

$P(WT = 5 \ min) = 0$

Definition. Consider a **probability density function**, pdf, of a random variable. It has the following properties:

(a) $f(x) \geq 0 \ \forall x \in \Omega$

(b) $\int_\Omega f(x)dx = 1$

In practical terms, $P(x \in A) = \int_A f(x)dx$, where A is the interval of interest.

Definition. By analogy, **cumulative distribution function** (cdf):

$$F(x) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

Example.

Sometimes we are concerned with the **median** *of the distribution. By definition, x_m is the median of a distribution if $P(X \leq x_m) = P(X \geq x_m) = \frac{1}{2}$. Using the cumulative function definition, it should be clear that*

$$F(x_m) = \int\limits_{-\infty}^{x_m} f(x)dx = \frac{1}{2}$$

Expected values

Intuition:

1. Sometimes instead of the distribution of a random variable we might be interested in a single number, a characteristic of its distribution. An obvious candidates is some "average" value of the distribution.
2. It turns out that this average value, the **mean** , is constant for each random variable with a given distribution

- For the **discrete** random variable x with the known probability distribution function, its mean is calculated as
$$E(x) = M(x) = \sum_i x_i P(x_i)$$
- For the **continuous** random variabe, the mean is
$$E(x) = \int_{\Omega} xf(x)dx$$

Example. *If you throw one die, what is the mean value?*

x	1	2	3	4	5	6
$P(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Solution.

$$E(x) = \bar{x} = \frac{1}{6} \cdot \sum_{i=1}^{6} i = \frac{21}{6} = \frac{7}{3}$$

Example. *What is the mean value if your throw two dice?*

x	2	3	4	...	12
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	...	$\frac{1}{36}$

Solution.

$$E(x) = \bar{x} = \sum_{i=2}^{12} x_i \cdot P(x) = \frac{2 \cdot 1}{36} + \frac{3 \cdot 2}{36} + \ldots = 7$$

Example. *The random variable x is distributed with the following distribution function:*

$$f(x) = \begin{cases} \frac{3}{8}x^2, & \textit{if} \quad 0 \leq x < 2; \\ 0, & \textit{elsewhere} \end{cases}$$

Solution.

$$E(x) = \int\limits_{0}^{2} xf(x)dx = \int\limits_{0}^{2} \frac{3}{8}x^3 dx = \frac{3}{8} \cdot \frac{x^4}{4}\bigg|_0^2 = \frac{3}{2}$$

Variation

The mean alone may be misleading when we compare two random variables with different distributions. Compare the following two random variables:

x	-0.01	0.01		y	-100	100
$P(x)$	0.5	0.5		$P(y)$	0.5	0.5

Both x and y have the same mean: $E(x) = E(y) = 0$, however it is important to account for the **deviation** of each variable from its mean.

Definition. Let's consider a new variable: **Deviation** from the mean: $x_i - E(x)$. For the random variable x, $E(x)$ is given, but x_i is random. Hence $x_i - E(x)$ is also a *random* variable. If we know the probability

$\Rightarrow$ P(x_i-E(x))=p_i if pmf is p_i=f(x)
\# E(x-E(x))=0
$\Rightarrow$ E(x)-E(E(x))=E(x)-E(x)=0
$\Rightarrow$ variation:
Var(x) $=\sum(x_i\text{-E(x)})^2\cdot f(x_i)$ discrete
$=\int_S (x\text{-E(x)})^2 f(x)dx$ cont.