Basic Probability

Prof. Sergio Focardi PhD
Email: sergio.focardi@edhec.edu

SECTION 1:

Outcome, Events,Probability

Mathematical representation of uncertainty

- Probability theory offers a mathematical representation of uncertainty
- Probability theory is a purely mathematical theory that was formalized by Kolmogorov in 1930
- How to interpret probability theory and how to connect probability concepts to experience is a subject of debate
- The two most widely used interpretations are the frequentist interpretation and the subjective interpretation (See slides 10-12)

Mathematical representation of uncertainty, ctd…

- We will first give a simplified presentation of probability concepts and then we will sketch the formal framework for the theory of probability
- The theory of probability is based on three main ingredients: outcomes, events, and probabilities
- In a nutshell, outcomes are possible observations, events are sets of observations, and probabilities are numbers between zero and one assigned to events

Experiments, outcomes, and the sample space

- A random experiment describes a process whose outcome is not known in advance with certainty but where all possible outcomes are known a priori.
- Example:
- A classical example is the rolling of a die where the possible outcomes are 1,2,3,4,5,6.

Definition

- A sample space is the collection of all possible outcomes of an experiment.
- Examples
- Rolling of a die: $S = \{1,2,3,4,5,6\}$.
- Outcome of two football games

Events

- Probabilities are not associated to individual outcomes but to events
- Events are subsets of the sample space, that is, they are sets of outcomes

Example of Events

The experiment consists of rolling a die once.

- The sample space S is $\{1,2,3,4,5,6\}$
- Define event A as 'score is lower than 4'. $A = \{1,2,3\}$
- Define event B as 'score is even'. $B = \{2,4,6\}$
- Define event C as 'score is 7' $\varnothing$

The experiment consists in observing the return of a financial investment

- The sample space is formed by all real numbers
- Event A: the investment earns a return of 10% or more.
- Event B: the investment earns a return below 10%.

Discrete and continuous sample spaces

- If the sample space is a discrete denumerable set we might assign a probability to events that include only one outcome
- However in continuous probability schemes the probability of individual outcomes is generally zero and probabilities are assigned only to events formed by infinite outcomes
- For example, if we assume that stock returns are continuous, the probability that a return r takes any given number, for example $r=0.05$, is zero
- We must therefore consider the probability that returns are in a given range, for example $0.02<r<0.05$

Probability

- The Probability P of an event E, denoted $P(E)$, is a measure of the likelihood that the event occurs
- P is a set function defined on the events of (subsets of) a sample space S and satisfies:

$$\forall s \in S$$

$$P(s) \geq 0$$

$$P(S) = 1$$

For all N pairwise disjoint events E_i and, when appropriate, for all infinite unions of pairwise disjoint events:

$$P\left(\bigcup E_i\right) = \sum_i P(E_i)$$

Interpretation of probabilities: the frequentist interpretation

- According to the frequentist interpretation we interpret the probability of an event as the relative frequency of the event in a large number of observations, that is, as the proportion of times an event happen in a large number of observations
- For example, the probability of each face of a die is 1/6; we interpret this fact through the assertion that in a large number of experiments each given face will appear approximately 1/6th of the times.
- For example, if we launch a die 12000 times we expect that the face 2 will appear approximately 2000 times

Interpretation of probabilities, ctd…

- However, it is problematic to make this statement precise from the theoretical point of view because there is no certainty, in a given sample, that relative frequency approximates probability
- Theoretically we can only say that in a large number of experiments the probability to observe a significant deviation of the relative frequency from the theoretical probability is very small
- There is no way to jump from probability to certainty
- Ultimately, in order to make this principle applicable we have to use some assumptio that rules out very unlikely events

Interpretation of probabilities: the judgmental interpretation

- According to the judgmental interpretation of probabilities we interpret probabilities as our intuitive judgment of the likelihood of an event
- Judgment of probability can be justified with theoretical motivations
- A judgment of probability can also be partially modified by empirical data. We start with a judgment of probability and we partially modify it through empirical data
- How to combine data and judgment is studied by Bayesian probability theory

σ-Algebras of events (this and the following two slides are more advanced)

- Thus far we have defined events simply as subsets of the sampling space
- However, the formal framework of probability theory first defines the class of admissible events. The class or collection of admissible events is called an algebra if finite, or a σ-algebra if infinite
- Given a sample space *U*, an algebra of events *F* is a collection of events such that, given any number *n* of events $A_1, A_2, \ldots, A_n$ that belong to *F* , the union $A_1 \cup \cdots \cup A_n$ and the intersection $A_1 \cap \cdots \cap A_n$ also belong to *F*
- In addition, the sample space belongs to *F* and, the empty set belongs to *F* and, if A is in *F* then the complement of A is also in *F*
- These conditions are redundant but it is useful to state them all explicitly
- A sigma algebra generalizes to infinite unions and intersections

Why algebras of events?

- For any sampling space there are many algebras of events: the smallest includes only the sampling space and the empty set, the largest includes all subsets of the sampling space
- Each algebra prescribes what sets we can recognize and therefore what information we have or we need
- A simple illustration is provided by the roulette
- The roulette is a game of chance performed in every casino
- It is formed by a disc with 37 (or 38) slots numbered from 0 to 36. A rolling ball randomly selects slots
- Zero is reserved while all other slots are available to gamblers

Why algebras of events? Ctd…

- Gamblers can bet on different schemes, for example single numbers or odd or even numbers
- If a gambler chooses to bet on single numbers then the appropriate algebra of events to analyze the game is that formed by all possible events
- However, if a gambler wants to bet only on odd/even numbers, then the appropriate algebra includes all numbers, all numbers less zero, no number, zero, even numbers, odd numbers
- Exercise: show the above

The probability space

- In summary, probabilities are defined over "probability spaces"
- A probability space is a triple (S,F,P) formed by:
- A sample space S of all possible outcomes of an experiment
- An algebra or a σ-algebra of events F
- A set function defined on events, called probability P, that assumes values from 0 and 1 that satisfies conditions in slide 9

Where you will find these concepts

- Probability is the usual theoretical tool to measure uncertainty
- You will find probabilities every time there is uncertainty and we want to quantify uncertainty and reason about uncertainty
- This includes asset management, risk management, corporate finance, economics

SECTION 2:

RULES OF PROBABILITY AND BAYES' THEOREM

Properties of probabilities

- The conditional probability of B given A is defined as:

$$P(B|A)=\frac{P(A\cap B)}{P(A)}$$

- The following rules hold:

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$
$$P(A^{C})=1-P(A)$$
$$P(A\cap B)=P(A)P(B|A)$$

Unconditional and conditional probabilities

- Provide answers to different questions
- Unconditional (marginal) probabilities $P(A)$ are the probabilities attached to the event A occurring.
- Conditional probabilities $P(A \mid B)$ are the probabilities attached to the event A occurring given (been known) that the event B has occurred.
- Example:
- What is the probability that a stock earns a return above the risk-free rate (event A) given that the stock earns a positive return (event B).

Independent events

- Two events A and B are independent if and only if their conditional probabilities equal their unconditional probabilities. $P(A \mid B) = P(A)$ or, equivalently $P(B \mid A) = P(B)$
- Multiplication Rule for Independent Events

$$P(A \cap B) = P(A)P(B)$$

Total probability rule

- We can analyze the likelihood of an event in the context of various scenarios.
- For each scenario (event) S we can define the scenario not-S called the complement of S and denoted S^C.

  $P(S^C) + P(S) = 1$
- Extension to n mutually exclusive and exhaustive scenarios

  $P(A) = P(A|S1)P(S1) + \cdots + P(A|SN)P(SN)$
- where $S_1, S_2, \ldots, S_N$ are mutually exclusive and exhaustive scenarios.

Bayes' theorem

- Bayes' theorem, despite its intrinsic simplicity, is arguably one of the most influential theorems of probability theory…
- because it originated an entire new view of statistics, Bayesian statistics.
- We will illustrate Bayes' Theorem as an elementary theorem to calculate inverse probabilities
- But Bayesian methods can be generalized to full-fledged Bayesian statistics and reasoning

Bayes' theorem

- As a theorem of elementary probability, Bayes' theorem is used to compute inverse probabilities. Recall Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- The demonstration is based on writing the joint probability in two different ways:

$$P(A \cap B) = P(B|A)P(A)$$
$$P(A \cap B) = P(A|B)P(B)$$
$$\Downarrow$$
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Example 1.

- The simplest context for inverse probabilities can be illustrated by the following example.

- Suppose there are two bowls on a table. Bowl A contains 5 red balls and 5 green balls and bowl B contains 8 red balls and 2 green balls. You choose at random one ball from one bowl. The chosen ball is red. What is the probability that the ball has been chosen from bowl A?

- The solution of this problem requires Bayes' theorem. In fact, we can easily compute direct probabilities and Bayes' theorem provides inverse probabilities.

Example 1., ctd…

- The probability of choosing the bowl A or B is 50% by the design of the experiment which suppose to choose at random between the two bowls.
-
- $P(A)=0.5$
- $P(B)=0.5$

Example 1., ctd…

- The probabilities of choosing red or green balls from each of the bowls are conditional probabilities
- Based on the assumption that balls are randomly chosen and given that the number of balls in each bowl is known, we write probabilities as follows:

- $P(red|A)=0.5$
- $P(green|A)=0.5$
- $P(red|B)=0.8$
- $P(green|B)=0.2$

Example 1., ctd…

- The unconditional probability of choosing a red or a green ball can be computed using the law of total probability. In fact:

$$P(\text{red}) = P(\text{red}|\text{A}) \times P(\text{A}) + P(\text{red}|\text{B}) \times P(\text{B}) = 0.5 \times 0.5 + 0.8 \times 0.5 = 0.65$$

$$P(\text{green}) = P(\text{green}|\text{A}) \times P(\text{A}) + P(\text{green}|\text{B}) \times P(\text{B}) = 0.5 \times 0.5 + 0.2 \times 0.5 = 0.35$$

- These probabilities are direct probabilities and are easy to compute. However, we want the reverse probabilities
- $P(A | \text{red})$
- $P(B | \text{red})$
- $P(A | \text{green})$
- $P(B | \text{green})$

Example 1., ctd...

- Bayes' theorem gives us the answer. In fact we can write:

$$P(\mathrm{A}|\mathrm{red}) = \frac{P(\mathrm{red}|\mathrm{A}) \times P(\mathrm{A})}{P(\mathrm{red})} = \frac{0.5 \times 0.5}{0.65} = 0.3846$$

$$P(\mathrm{B}|\mathrm{red}) = \frac{P(\mathrm{red}|\mathrm{B}) \times P(\mathrm{B})}{P(\mathrm{red})} = \frac{0.8 \times 0.5}{0.65} = 0.6153$$

- The two probabilities obviously sum up to 1 (actually to 0.9999 due to rounding). We can apply the same reasoning if we have chosen a green ball

Bayes' theorem and information

- The above example is an example of Bayes' theorem applied to classical statistics.
- We can interpret Bayes' theorem as a method to move from unconditional probabilities $P(A), P(B)$ to conditional probabilities such as $P(A \mid \text{red})$.
- In this interpretation, we add information in the classical framework of conditional probabilities.
- In this interpretation, let's consider an additional example on the practice of testing people or artifacts.

A car test

- Consider an automatic test to detect the good functioning of a car engine at a car check-up shop.
- If an engine passes the test it is considered fit, if not the car must be sent to a repair shop.
- The test is not perfect. The test might fail to detect a faulty engine.
- And under certain conditions it might reject a perfectly functioning engine.

Sensitive and specific

- Suppose the relevant probabilities are the following.
- The test is 99% sensitive, that is it correctly rejects faulty engines in 99% of the cases
- The test is 99% specific, that is it correctly identify good engines in 99% of the cases

Direct probabilities…

- These are direct probabilities that can be estimated as relative frequencies
- Call *P* the event that an engine pass the test and *R* the event that the engine is rejected, call *G* the event that an engine is good and *F* the event that an engine is faulty
- The two previous probabilities can be written as:

  $$P(R|F) = 0.99,\ P(P|G) = 0.99$$

- That is, the probability of rejection given the engine is faulty is 0.99, the probability that a good engine passes the test is also 0.99

And inverse probabilities

- Now you take your car to the repair shop and it is rejected.
- What is the probability that the car is really faulty?
- You have to compute , $P(F|R)$ that is, probability that it is faulty given it is rejected.
- Bayes' theorem writes as follows:

$$P(F|R) = P(R|F)\frac{P(F)}{P(R)} = 0.99\frac{P(F)}{P(R)}$$

A surprising result

- The unconditional probability of rejection is:

$$P(\mathrm{R}) = P(\mathrm{R}|\mathrm{F})P(\mathrm{F}) + P(\mathrm{R}|\mathrm{G})P(\mathrm{G}) =$$
$$P(\mathrm{R}|\mathrm{F})P(\mathrm{F}) + [1 - P(\mathrm{P}|\mathrm{G})] \times [1 - P(\mathrm{F})] =$$
$$0.99 \times P(\mathrm{F}) + 0.01 \times [1 - P(\mathrm{F})]$$

- Therefore, the probability that your car is really faulty given the test results hinges on the unconditional probability that a generic car is faulty.
- Suppose that the probability that a car is faulty be 1%.

A surprising result, ctd...

- Then:

$$P(\mathrm{F}|\mathrm{R}) = 0.99\frac{P(\mathrm{F})}{0.99\times P(\mathrm{F})+0.01\times[1-P(\mathrm{F})]} = \frac{0.01}{0.99\times 0.01+0.01\times 0.99} = \frac{1}{2\times 0.99} = 0.5$$

- Suppose that the percentage of faulty cars is much higher, say 5%. The previous formula yields:

$$P(\mathrm{F}|\mathrm{R}) = 0.99\frac{P(\mathrm{F})}{0.99\times P(\mathrm{F})+0.01\times[1-P(\mathrm{F})]} = \frac{0.05}{0.99\times 0.05+0.01\times 0.95} == 0.8390$$

- Suppose now that the percentage of faulty cars is lower, say 0.5%. The previous formula yields:

$$P(\mathrm{F}|\mathrm{R}) = 0.99\frac{P(\mathrm{F})}{0.99\times P(\mathrm{F})+0.01\times[1-P(\mathrm{F})]} = \frac{0.005}{0.99\times 0.005+0.01\times 0.995} == 0.3322$$

To practice: a medical test

- You are a young person. After a routine medical check you are told that you have a rare but very dangerous disease
- You would like to go through a second medical test but the results of the new test will be known only in a month.
- Your physician wants you to start immediately a therapy which is painful and has serious side consequences.
- You are deeply worried and you begin to investigate different sources of information.
- You find what you consider a highly reliable medical website which gives you the following information:
- The disease you are supposed to have is very rare in people below 40 years of age. Actually the probability that a person below 40 years of age has this disease is 1/100,000

Medical test, ctd...

- The sensitivity of the test used for your diagnosis is 99.9% , that is, the probability that the disease is detected in a person affected by the disease is 99.9% (0.999)
- The specificity of that test is 99.9%, that is the probability that the disease is wrongly diagnosed in a healthy person is 0.1%.

- How do you use this information to reach a decision whether to start therapy immediately or to wait for a confirmation test?

Solution exercise

The probability that you have the disease given that the test is positive can be computed as follows. Call P(C) the unconditional probability of disease

$$P(\text{disease}) = 0.00001$$

$$P(\text{healthy}) = 0.99999$$

$$P(\text{positive}|\text{disease}) = 0.999$$

$$P(\text{positive}|\text{healthy}) = 0.999$$

$$P(\text{positive}) = P(\text{positive}|\text{disease}) \times P(\text{disease}) + P(\text{positive}|\text{healthy}) \times P(\text{healthy})$$

$$0.999 \times 0.00001 + 0.001 \times 0.99999 = 0.001$$

$$P(\text{disease}|\text{positive}) = \frac{P(\text{positive}|\text{disease}) \times P(\text{disease})}{P(\text{positive})} = \frac{0.999 \times 0.00001}{0.001} = 0.01$$

Mixing judgment and data

- When making investment decisions, we often start with some a priori viewpoint based on our previous experience and acquired knowledge.
- These a priori viewpoints may be changed by new knowledge or observations.
- The Bayes' formula is a rational way to adjust our viewpoints given some new information.
- The idea of the formula
- P(Event|Information) = P(Information|Event) P(Event)/P(Information)
- The Bayes' formula relates posterior with prior probabilities and can
- be used to update subjective beliefs.
- $P(A|B) = P(A) \cdot P(B|A)$
- $P(A)P(B|A)+P(Ac)P(B|Ac)$

Where you will find these concepts

- Rules of probability are used in every probabilistic context
- Bayesian analysis is used in many different contexts in econometrics
- You will find Bayesian theory especially in asset management as a way to mix data and models with personal views and judgment

SECTION 3:

RANDOM VARIABLES AND EXPECTATION

Random Variables

- Thus far we have discussed probability defined on sample spaces
- In practice, however, we work with observable quantities represented by random variables
- A random variable X is a variable whose value is not known but such that we can define the probability that the variable X falls in any given interval $a<X<b$
- The defining property of a random variable is the possibility to assign a probability to each interval $(a<X<b)$

Random variables, ctd…

- The intuition behind the concept of a random variable is based on the fact that we associate a probability to each interval
- For example, we can think of stock returns as random variables because we can define the probability that a stock returns fall in a given interval, eg $0.02<r<0.05$
- But how do we define probabilities associated to intervals of random variables?

Random variables and probability space (more advanced)

- In the previous slides we discussed how probabilities are defined over a probability space (S, F, P)
- We define a random variable X as a real valued function that projects (maps) any outcome s of an experiment to a real number: $X(s): S \rightarrow R$

  so that all sets $(a<X<b)$ for any real numbers a,b and plus/minus infinite are events
- The probability that $(a<X<b)$ is the probability of the event that corresponds to (a,b) that is $s: a < X(s) < b$

Do we need a probability space?

- The above is typical of probability theory as used in finance
- In finance and in economics we assume that there is a probability space formed by the "states of world"
- A state of the world is a complete possible history of a given economy, a possible scenario
- Random variables such as returns or prices are functions defined over the set of possible states of the world
- Many states of the world can share the same value of a variable

Do we need a probability space? Ctd…

- Is there a way to define probabilities for random variables without assuming an underlying probability space?
- The answer is yes, there is a different probability theory called Algebraic Probability Theory, introduced in the 1930s by Von Neumann, which is based directly on random variables
- Algebraic probability theory is a generalization of classical probability theory but it is not simpler and it is rarely used in finance

Discrete random variables

- Definition:
- Discrete random variables are those that have only a discrete set of possible outcomes
- That is, they can assume only a finite or denumerably infinite set of values
- Example: stock prices quoted in tenth of dollars (deumerably infinite)
- Example: the labels 1,2,…,6 of the faces of a die (finite)

Continuous Random Variables

- Continuous random variables are those which relate to measuring quantities such as time and asset returns
- The number of possible outcomes is uncountably infinite
- Then, it only makes sense to assign probabilities to the event that a continuous random variable takes values between two limits (not that it takes a specific value)

- Example: stock prices can be thought as continuous variables because they can assume any value

Properties

- A constant is a random variable
- The product aX is a random variable
- The sum of random variables $X+Y$ is a random variable
- The product of random variables XY is a random variable

Cumulative Distribution Functions-CDF

- Definition
- The cumulative distribution function $F(x)$of a random variable X is defined as
- $F(x) = P(X \leq x)$
- TO NOTE: X is a random variable, x is a real valued variable, $F(x)$ is a real valued function of a real valued variable x not of the random variable X
- The following properties of F hold

$$0 \leq F(x) \leq 1$$

$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to +\infty} F(x) = 1$$

Cumulative Distribution Functions-CDF, ctd…

- The CDF $F(x)$ of a random variable X always exists because by definition $P(X<x)$ is defined for any x
- Both continuous and discrete variables admit a CDF
- The CDF of a discrete random variable is a step function
- The CDF of a continuous random variable is generally a continuous function (discontinuities might exist only if a non zero probability is concentrated in single values)

Probability function

- If X is discrete, we can define the probability of each value and the probability function which assigns to each possible value its probability:

$$p_i = P\left(X = x_i\right)$$

$$\sum_i p_i = 1$$

Probability Density Functions-PDF

- If X is continuous and $F(x)$ is differentiable its derivative $f(x)=F'(x)$ is called the Probability Density Function – pdf – of X and the following relationship holds

$$F(x)=\int_{-\infty}^{x} f(u)du$$

- For a continuous random variable, X, its probability density function (pdf) , f (x) has the following properties

$$f(x)\geq 0,\ (-\infty < x < +\infty)$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

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

Expectation

- Intuitively, the expectation of a random variable is the sum of all its possible values weighted with their relative probabilities
- If the variable X is discrete then $E(X) = \sum p_i x_i$ where the sum is finite or infinite depending whether X assumes a finite or infinitely denumerable set of values x_i
- If the variable X is continuous and admits a pdf then:

$$E(X) = \int_{-\infty}^{+\infty} f(x) x dx$$

- The expectation of a random variable is also called the mean of the random variable

Moments

- The *nth* raw moment of a random variable is defined as:

$$\tilde{\mu}_n = E\left(X^n\right)$$

- The *nth* central moment $n>1$ of a random variable is defined as:

$$\mu_n = E\left[\left(X - \mu_X\right)^n\right]$$

- In particular, the variance is the second central moment:

$$\mu_1 = E\left(X\right)$$

$$\mu_2 = E\left[\left(X - \mu_1\right)^2\right]$$

Skewness and Kurtosis

- Skewness is defined as:

$$S = \frac{\mu_3}{\mu_2^{\frac{3}{2}}}$$

- Kurtosis is defined as:

$$K = \frac{\mu_4}{\mu_2^2}$$

Joint distributions

- Given two random variables X, Y
- The joint cumulative distribution $F(x,y)$ is defined as the probability that $X<x$ and $Y<y$
- $F(x,y)=P(X<x \text{ and } Y<y)$
- The joint density is defined by the relationship:

$$F(x,y) = \iint_{u \le x, v \le y} f(u,v)\,du\,dv$$

Conditional density

The conditional density is defined as:

$$f\left(x \middle| y\right) = \frac{f\left(x, y\right)}{f\left(y\right)}$$

Independent variables

The variables X, Y with joint pdf f are said to be independent if

$$f(x, y) = f(x)f(y)$$

Properties of expectation

- The expectation has the following properties:

$$E(b)=b$$

$$E(aX)=aE(X)$$

$$E(X+Y)=E(x)+E(y)$$

$$E(aX+b)=aE(x)+b$$

- However the property: $E(XY)=E(X)E(Y)$

  holds only if the variables are independent

Conditional expectation

We can compute the expectation Y given X

using the conditional density:

$$E(X|Y) = \int_{-\infty}^{+\infty} u \frac{f(u, y)}{f(y)} du$$

Variance and standard deviation

- The variance is defined as:

$$\mu_X = \overline{X} = E(X)$$

$$\operatorname{var}(X) = \sigma_X^2 = E\left[(X - \mu_X)^2\right] = \int_{-\infty}^{+\infty} (u - \mu_X)^2 f(u)du$$

- The standard deviation σ_X is the square root of the variance

Properties of the variance

- The following properties hold:

$$\operatorname{var}(X+b)=\operatorname{var}(X)$$

$$\operatorname{var}(aX)=a^2\operatorname{var}(X)$$

$$\operatorname{var}(aX+b)=a^2\operatorname{var}(X)$$

- But $\operatorname{var}(X+Y)=\operatorname{var}(X)+\operatorname{var}(Y)$

  only if X,Y independent

Covariance

The Covariance between two random variables X and Y is defined as follows:

$$\mu_X = E(X)$$

$$\mu_Y = E(Y)$$

$$\operatorname{cov}(X,Y) = \sigma_{XY} = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$

Properties of covariance

The following properties hold:

$$\operatorname{cov}(X,Y) = E(XY) - \mu_X \mu_Y$$

$$\operatorname{var}(X) \equiv \operatorname{cov}(X,X) = E(X^2) - \mu_X^2$$

$$\operatorname{cov}(aX+b, cY+d) = ac\operatorname{cov}(X,Y)$$

$$Y = V + W$$

$$\operatorname{cov}(X,Y) = \operatorname{cov}(X,V) + \operatorname{cov}(X,W)$$

Correlation coefficient

The correlation coefficient is defined as follows:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Binomial variable

- The binomial random variable X is the number of successes in n independent trials when the probability of success on each trial is a constant p.
- We write $X \sim B(n, p)$
- There are n+1 possible values for X, $0, 1, 2, \cdots, n$.

Binomial distribution

- The probability of r successes is:

$$P(X = r) = \binom{n}{r} p^r (1-p)^{n-r} = \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}, \ 0 \le r \le n$$

- The following properties hold:

$$E(X) = np$$

$$\mathrm{var}(X) = np(1-p)$$

$$\sigma_X = \sqrt{np(1-p)}$$

Uniform continuous distribution

Constant density in an interval a,b:

$$f(x) = \frac{1}{b-a},\ \ a \le x \le b$$

$$0,\ \ x < a, x > b$$

$$U \sim Uniform(a,b)$$

$$E(U) = \frac{a+b}{2}$$

$$\mathrm{var}(U) = \frac{(b-a)^2}{12}$$

Normal distribution

- A random variable X is said to be normally distributed

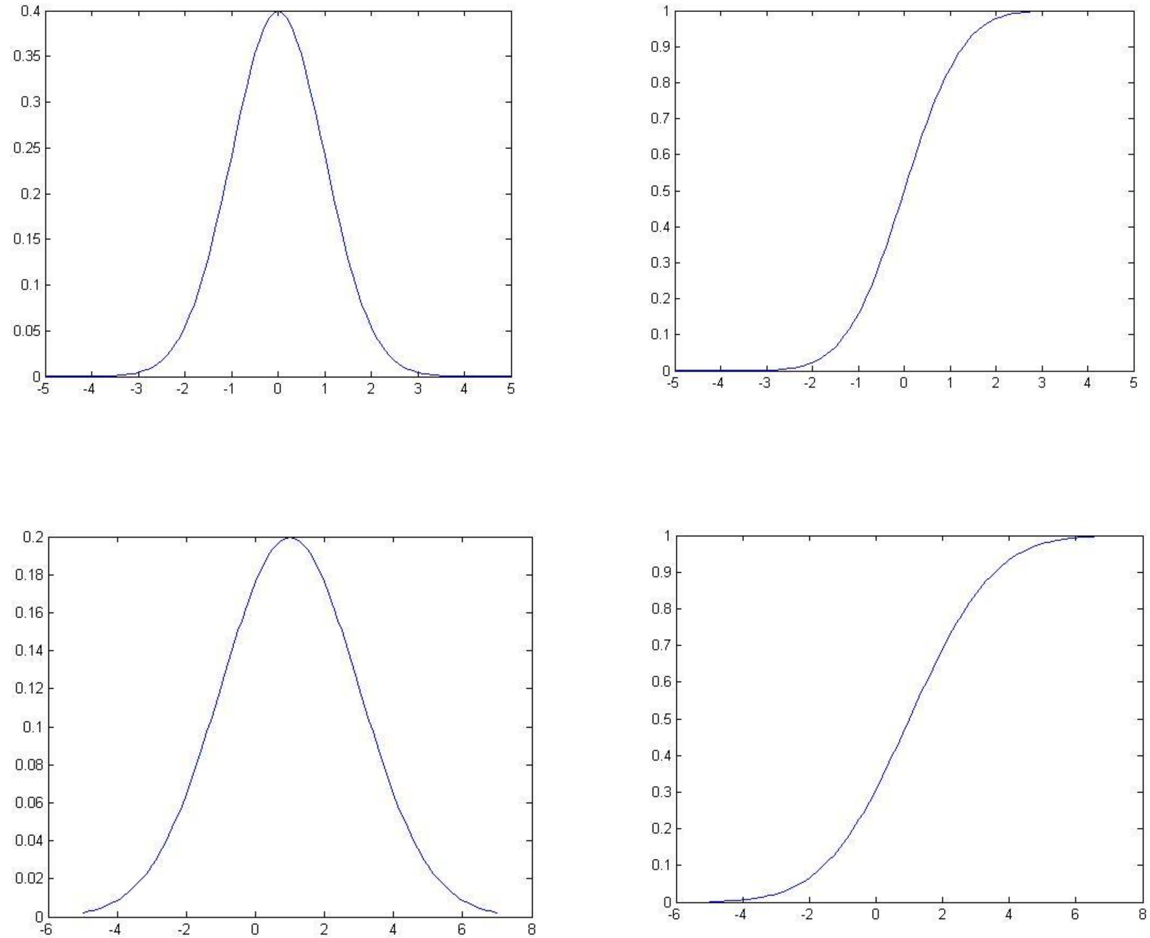
$$X \sim N(\sigma, \mu)$$

- If it has the following 2-parameter probability density function, where μ, σ are respectively the mean and the standard deviation of X

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

- Standard normal variable Z has mean zero and std=1
- We standardize by defining $Z = \dfrac{X - \mu}{\sigma}$

Normal variables: 0,1 and 1,2



Linear Regression

- Suppose that two random variables X, Y are approximately linked by a linear relationship $Y \approx a + bX$ where a, b are constants
- This means that when we observe a value x for the variable X we also observe or we can predict that we would observe a value y of the other variable approximately equal to $y \approx a + bx$
- We can make this concept precise in a probabilistic sense through the concept of linear regression: $Y \approx a + bX + \varepsilon$ where ε is a random variable

Where you will find these concepts

- Random variables are used when there is uncertainty as regards the value assumed by given quantities such as prices or returns
- Probability densities are commonly used in application of probability to financial and economic problems
- You will find probability density functions in most disciplines you will study, finance, corporate finance, economics

SECTION 4:
EXERCISES

Exercise

- Suppose you have two limit orders outstanding on two different stocks.
- The probability that the first limit order executes before the close of trading is 0.45.
- The probability that the second limit order executes before the close of trading is 0.20.
- The probability that the two orders both execute before the close of trading is 0.10.
- What is the probability that at least one of the two limit orders executes before the close of trading?

- A=(first limit order executes before the close of trading)
- B=(second limit order executes before the close of trading)

$$P(A)=0.45,\ \ P(B)=0.2,\ \ P(A\cap B)=0.1$$
$$P(A\cup B)=P(A)+P(B)-P(A\cap B)=0.45+0.2-0.1=0.55$$

Exercise

- Suppose that 5 percent of the stocks meeting your stock-selection criteria are in the telecommunications (telecom) industry.
- Dividend-paying telecom stocks are 1 percent of the total number of stocks meeting your selection criteria.
- What is the probability that a stock is dividend paying, given that it is a telecom stock that has met your stock selection criteria?

- A: (dividend paying)
- B: (telecom)
- C: (meet selection criteria)
- P(BC)=0.05
- P(ABC)=0.01

$$P(B \cap C) = 0.05,\ \ P(A \cap B \cap C) = P(A \cap (B \cap C)) = 0.01$$

$$P(A | (B \cap C)) = \frac{P(A \cap (B \cap C))}{P(B \cap C)} == \frac{0.01}{0.05} = 0.2$$

Exercise

- You are using the following three criteria to screen potential acquisition targets from a list of 500 companies:

Criterion	Fraction of the 500 Companies Meeting the Criterion
Product lines compatible	0.20
Company will increase combined sales growth rate	0.45
Balance sheet impact manageable	0.78

- If the criteria are independent, how many companies will pass the screen?

$0.2 x 0.45 x 0.78 = 0.0702$,

$0.0702 x 500 = 35$

Exercise

- 15. You have developed a set of criteria for evaluating distressed credits.
- Companies that do not receive a passing score are classed as likely to go bankrupt within 12 months.
- You gathered the following information when validating the criteria:
- Forty percent of the companies to which the test is administered will go bankrupt within 12 months: $P(nonsurvivor) = 0.40$.
- Fifty-five percent of the companies to which the test is administered pass it: $P(pass\ test) = 0.55$.
- The probability that a company will pass the test given that it will subsequently survive 12 months, is 0.85: $P(pass\ test \mid survivor) = 0.85$.
- A. What is $P(pass\ test \mid nonsurvivor)$?
- B. Using Bayes' formula, calculate the probability that a company is a survivor, given that it passes the test; that is, calculate $P(survivor \mid pass\ test)$.
- C. What is the probability that a company is a *nonsurvivor*, given that it fails the test?
- D. Is the test effective?

A)

$P(\text{Pass test}) = P(\text{Pass test} \mid \text{survivor})P(\text{survivor})) + P(\text{Pass test} \mid \text{nonsurvivor})P(\text{nonsurvivor}))$

$P(\text{survivor}) = 1 - P(\text{nonsurvivor}) = 1 - 0.4 = 0.6$

$P(\text{Pass test}) = 0.55 = 0.85 \text{x} 0.6 + P(\text{Pass test} \mid \text{nonsurvivor}) \text{x} 0.4$

$P(\text{Pass test} \mid \text{nonsurvivor}) = (0.55 - 0.85 \text{x} 0.6)/0.4 = 0.1$

B)

$P(\text{Survivor} \mid \text{pass test}) = P(\text{Pass test} \mid \text{survivor})P(\text{survivor}) / P(\text{Pass test})$

Test on longevity

- In a given population the probability that a person born today lives to the age of 40 is 0.98 and the probability that a person born today lives to the age of 80 is 0.85.
- What is the probability that a person born today reaches the age of 80 given that he/she has reached 40?

- A: lives up to 40
- B: lives up to 80

$$
\begin{aligned}
&P(A) = 0.98 \\
&P(B) = 0.85 \\
&P(A \cap B) = P(B) = 0.85 \\
&P(B|A) = \frac{P(A \cap B) \times P(B)}{P(A)} = \frac{0.85}{0.98} = 0.8673
\end{aligned}
$$

- if

$$
\begin{aligned}
&P(A) = 0.75 \\
&P(B) = 0.7
\end{aligned}
$$

- then

$$
\begin{aligned}
&P(A \cap B) = P(B) = 0.7 \\
&P(B|A) = \frac{P(A \cap B) \times P(B)}{P(A)} = \frac{0.7}{0.75} = 0.933
\end{aligned}
$$