Probabilities
Conditional Probabilities
Random Experiments and Variables

Sorana D. Bolboacă

Abraham de Moivre (1667-1754)

"Some of the Problems about Chance having a great appearance of Simplicity, the Mind is easily drawn into a belief, that their Solution may be attained by the meer Strength of natural good Sense; which generally proving otherwise and the Mistakes occasioned thereby being not infrequent, 'tis presumed that a Book of this Kind, which teaches to distinguish Truth from what seems so nearly to resemble it, will be looked upon as a help to good Reasoning."

Objectives

- Basics of probabilities. Events space.
- Probabilities rules
- Conditional probabilities. Risks and rates
- Random experiment and random variables

Basics of Probabilities

- Probability
  - a numerical measure of uncertainty
    - Uncertainty about phenomena or event outcomes could be described in terms such as “unlikely”, “possible”, “likely”, “probable”
  - a way of expressing knowledge or belief that an event will occur or has occurred
  - is synonymous with chance or likelihood
  - Is a numerical value of uncertainty associated with the outcome of an uncertain or unpredictable event

Basics of Probabilities

- Experiment:
  - = an activity or investigation for which the results are uncertain
    - Tossing a coin
    - Rolling a die
    - Counting failures over time
- Outcome:
  - = the result of one execution of the experiment
- Events space = the set of all possible outcomes of an experiment.
  - Symbol: S

BASICS OF PROBABILITIES

- Event spaces:
  - Tossing a coin:
    - Two possible outcomes: head (H) and tails (T)
    - S = {H, T}
  - Rolling a die:
    - Six possible outcomes: 1, 2, 3, 4, 5, 6
    - S = {1, 2, 3, 4, 5, 6}
- The event spaces can contain discrete points (such as pass or fail) or continuum points. An event E is a specified set of possible outcomes in a event space S: $E \subset S$, where $\subset$ = "subset of"

BASICS OF PROBABILITIES

- Empty set of null set: a special set without any element. It is denoted by $\emptyset$.
- Elementary event: the event that could not be decomposed into smaller parts.
- Compound event: the event that could be decomposed into smaller parts.
  - In practical situations most events are compound events

BASICS OF PROBABILITIES

- **Reunion (OR):**
  - Symbol: A∪B
    - At least one event (A OR B) occurs
- **Intersection (AND)**:
  - The probability that A and B both occur
  - Use the multiplication rule
  - Symbol: A∩B
    - the events A and B occur simultaneously
- **Negation:**
  - Symbol: nonA

Basic of Probabilities

- Complementary event:
  - A complement of an event E in a event space S is the set of all event points in S that are not in E.
  - Symbol: $\overline{E}$ (read E-bar)
- Mutually exclusive events:
  - Two events E_1 and E_2 in a event space S are said to be mutually exclusive if the event $E_1 \cap E_2$ contain no outcome in the event space S.
  - These two events can not happen in the same time.

BASICS OF PROBABILITIES

- Independent events:
  - Two events E_1 and E_2 are said to be statistically independent events if the joint probability of E_1 and E_2 equals the product of their marginal probabilities of E_1 and E_2.
- Union: The union of two events E_1 and E_2, denoted by $E_1 \cup E_2$, is defined to be the event that contain all sample points in E_1 or E_2 or both.
- Intersection: The intersection of two events E_1 and E_2, denoted by $E_1 \cap E_2$, is defined to be the event that contain all sample points that are in both E_1 and E_2.

BASICS OF PROBABILITIES

- Marginal probability: the probability that an event will occur, regardless of whether other events occurs
  - The probability Pr(R) of a car jumping the red line at a given time at a given intersection, regardless of whether a pedestrian is in the crosswalk
- Conditional probability: of event B, given that event A has occurred, denoted by Pr(B|A), is defined by:

$$\Pr(B|A) = \Pr(B \cap A)/\Pr(B) \text{ for } \Pr(B)>0$$

Basics of Probabilities

Subjective probability:

- Established subjective (empiric) base on previous experience or on studying large populations
- Implies elementary that are not equipossible (equally likely)

Objective probability:

- Equiprobable outcomes
- Geometric probability

Formula of calculus:

- If an A event could be obtained in S tests out of n equiprobable tests, then the Pr(A) is given by the number of possible cases
- Pr(A) = (no of favorable cases)/(no of possible cases)

Basics of Probabilities

- A probability of an event A is represented by a real number in the range from 0 to 1 and written as Pr(A).
- Probabilities are numbers which describe the likelihoods of random events.

$$\mathbf{Pr(A) \in [0, 1]}$$

- Let A be an event:
  - Pr(A) = the probability of event A
  - If A is certain, then Pr(A) = 1
  - If A is impossible, then Pr(A) = 0

CHANCES AND ODDS

- Chances are probabilities expressed as percent.
  - Range from 0% to 100%.
  - Ex: a probability of 0.65 is the same as a 65% chance.

- The odds for an event is the probability that the event happens, divided by the probability that the event doesn't happen.
  - Can take any positive value
  - Let A be the event. Odds(A) = Pr(A)/[1-Pr(A)], where 1-Pr(A) = Pr(nonA)
  - Example: Pr(A) = 0.75; a probability of 0.75 is the same as 3-to-1 odds (0.75/(1-0.75)=0.75/0.25=3/1)

Properties of Probabilities

- Take values between 0 and 1:

$$0 \le Pr(A) \le 1$$

- Pr(event space) = 1
- The probability that something happens is one minus the probability that it does not:

$$Pr(A) = 1 - Pr(nonA)$$

- **Addition Rule**: probability of A **or** B:

$$Pr(A \cup B) = Pr(A) + Pr(B)$$

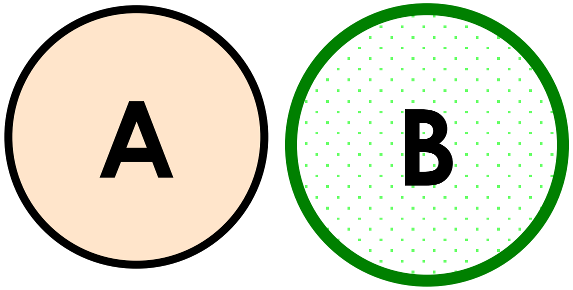
when A and B are mutually exclusive

- **Multiplication Rule**: probability of A **and** B:

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

when A and B are independent

PROBABILITY RULES: ADDITION RULE



- Let A and B be two events:

$$\mathbf{Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)}$$

$$\mathbf{Pr(A\ or\ B) = Pr(A) + Pr(B) - Pr(A\ and\ B)}$$

- A and B mutually exclusive:
  - **$Pr(A \cap B) = 0$**
  - **Pr(A and B) = 0**
- A = {SBP of mother > 140 mmHg}
  - Pr(A) = 0.25
- B = {SBP of father > 140 mmHg}
  - Pr(B) = 0.15
- What is the probability that mother or father to have hypertension?

$$\mathbf{Pr(A \cup B) = 0.25 + 0.15 - 0 = 0.40}$$

$$\mathbf{Pr(A\ or\ B) = 0.25 + 0.15 - 0 = 0.40}$$

PROBABILITY RULES: ADDITION RULE

- In a cafe are at a moment 20 people, 10 like tea, 10 like coffee and 2 like tea and coffee.
- What is probability to random extract from this population one person who like tea or coffee?

$$\mathbf{Pr(tea \cup coffee) = Pr(tea) + Pr(coffee) - Pr(tea \cap coffee)}$$

$$\mathbf{Pr(tea\ or\ coffee) = Pr(tea) + Pr(coffee) - Pr(tea\ and\ coffee)}$$

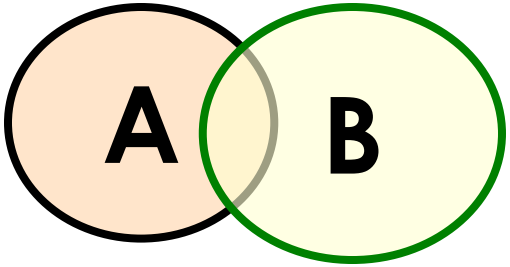
$$\mathbf{Pr(tea \cup coffee) = 0.50 + 0.50 – 0.10 = 0.90}$$

PROBABILITY RULES: MULTIPLICATION RULE

- Let A and B be two events:

$$\mathbf{Pr(A \cap B) = Pr(A) \cdot Pr(B|A)}$$

$$\mathbf{Pr(A\ and\ B) = Pr(A) \cdot Pr(B|A)}$$



- Independent events: **Pr(B|A) = Pr(B)**

- A = {SBP of mother > 140 mmHg}
  - Pr(A) = 0.10
- B = {SBP of father > 140 mmHg}
  - Pr(B) = 0.20
- $Pr(A \cap B) = 0.05$; Pr(A and B) = 0.05
- The two events are dependent or independent?

$$\mathbf{Pr(A \cap B) = Pr(A) \cdot Pr(B)} \text{ – \textbf{independent events}}$$

$$0.05 \neq 0.10 * 0.20 \rightarrow \text{the events are dependent}$$

Bayes' Theorem

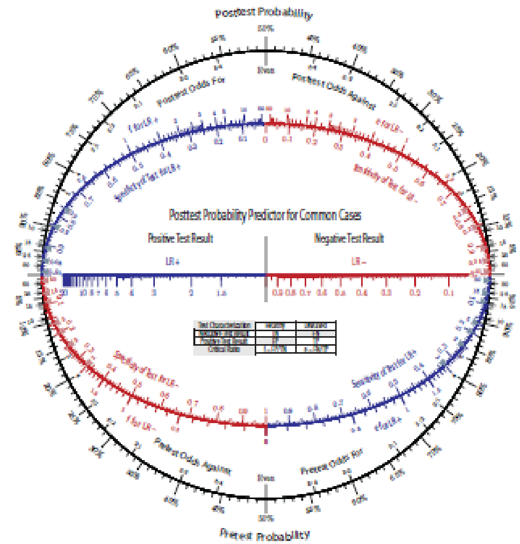
- Bayes' law / Bayes' rule
- Gives the relationship between the probabilities if event A and event B and the conditional probabilities of A given B and B given A:

$$\Pr(A|B) = \frac{\Pr(B|A) \times \Pr(A)}{\Pr(B|A) \times \Pr(A) + \Pr(B|nonA) \times \Pr(nonA)}$$

BAYES' THEOREM

Diagnosis

It's all about probabilities



Bayer's rule → Nomography (Modern)

Math for updating probabilities

Pr_{old} + Test = Pr_{new}

Graphical technique for doing the math

Bayesian Theorem

- Tests are not events
- Tests detect things that do not exists (false positive) and could miss thinks that exists (false negative)
- Tests give us test probabilities NOT the real probability
- False positive skew results
- People prefer natural numbers as 100 in 10.000 rather than 1%
- Even science is a test

Bayesian Theorem

- Bayes' theorem finds the actual probability of an event from the results of a test
  - Correct the measurement errors
    - If you know the real probability & the chance of false positive and false negative
  - Relate the actual probability to the measured test probability.
    - Given mammogram test results and known error rates, you can predict the actual chance of having cancer

Anatomy of a test

- 2% of women have breast cancer
- 70% of mammograms detect breast cancer when it is present → 30% of mammograms miss the diagnosis
- 10% of mammograms incorrectly detect breast cancer when it is **not** present → 90% of mammograms correctly return a negative result

	Breast cancer + (2%)	Breast cancer - (98%)
Mammo +	70	10
Mammo-	30	90

Anatomy of a test: how to read it!

- 2% of women have breast cancer
- If you already have breast cancer, there is 70% chance that your mammogram will be positive and 30% chance that your test will to be negative
- If you do not have breast cancer, there is 10% chance that your mammogram will be positive and 90% chance that your test will to be negative

	Breast cancer + (2%)	Breast cancer - (98%)
Mammo +	70	10
Mammo-	30	90

Anatomy of a test

- Suppose that you have a patient with a positive result. What are her chances to have breast cancer?

	Breast cancer + (2%)	Breast cancer - (98%)
Mammo +	70	10
Mammo-	30	90

- The probability of a **true positive** = the probability to have breast cancer × probability that the mammogram to be positive = 0.02 × 0.7 = 0.014 → 1.4% chance
- The probability of a **false positive** = the probability not to have breast cancer × probability that the mammogram to be positive = 0. 98 × 0.10 = 0.098 → 9.8% chance

Anatomy of a test

	Breast cancer + (2%)	Breast cancer - (98%)
Mammo +	True positive 0.02*0.70 = 0.014	False positive 0.10*0.98=0.098
Mammo-	False negative 0.02*30 = 0.600	True negative 0.90*0.98 = 0.882

- What is the chance that your patient to really have cancer if she get a positive mammogram.
  - The chance of an event is the number of ways it could happen given all possible outcomes:
  - Probability = (desired event)/ (all possibilities)
  - Probability = 0.014/(0.014+0.098) = 0.125 → the chance that your patient to have breast cancer if the mammogram is positive is 12.5%

Anatomy of a test

- → a positive mammogram only means that the individual chance of breast cancer is 12.5%, rather than expected 70%
- → the mammogram gives a false positive 10% of the time → there will be false positives in any given population
- → the problem can be turned into an equation = Bayes' Theorem

Bayes' Theorem

$$\Pr(A \mid B) = \frac{\Pr(B \mid A) \times P(A)}{\Pr(B \mid A) \times \Pr(A) + \Pr(B \mid nonA) \times \Pr(nonA)}$$

$$\Pr(A \mid B) = \frac{\Pr(B \mid A) \times P(A)}{\Pr(B)}$$

- P(A|B) = probability of having breast cancer (A) given a positive mammogram (B)
- P(B|A) = probability of a positive mammogram (B) given that breast cancer is present (A)
- P(A) = probability of having breast cancer → P(nonA) = probability not to have cancer
- P(B|nonA) = probability of a positive mammogram

Prior vs posterior probability

- Prior probability
  - what you believe before seeing any data
- Posterior probability
  - P(hypothesis|data) = the probability of a hypothesis given the data
  - depends on prior probability & observed data

Bayesian inference

- Prior probability + data → posterior probability

Published scientific literature / National statistical data

- P(hypothesis is true | observed data)

A good prior is helps, a bad prior hurts, but the prior matter less the more data you have!

CONDITIONAL PROBABILITIES

- Conditional probability: example
  - (Tuberculin Test+|TBC) is the probability of obtaining a positive tuberculin test to a patient with tuberculosis
- **Pr(B|A)** is not the same things as **Pr(A|B)**

	TBC+	TBC-
Test+	15	12
Test-	25	18

- Let:
  - A = {TBC+}
  - B = {Tuberculin Test+}
- Pr(A) = (15+25)/(15+12+25+18) = 0.57 (prevalence)
- Pr(nonA) = (12+18)/(15+12+25+18) = 0.43
- Pr(B|A) = probability of a positive tuberculin test to a patient with TBC = 15/(15+25) = 0.38 = **Sensibility (Se)**

CONDITIONAL PROBABILITIES

- Let:
  - A = {TBC+}
  - B = {Tuberculin Test+}

	TBC+	TBC-
Test+	15	12
Test-	25	18

- Pr(nonB|nonA) = probability of obtaining a negative test to a patient without TBC = 18/(18+12) = 0.60 = **Specificity (Sp)**
- Pr(A|B) = probability that a person with TBC to have a positive tuberculin test = 15/(15+12) = 0.56 = Predictive Positive Value **(PPV)**

CONDITIONAL PROBABILITIES

- Let:
  - A = {TBC+}
  - B = {Tuberculin Test+}

	TBC+	TBC-
Test+	15	12
Test-	25	18

- Pr(nonA|nonB) = probability that a person without TBC to have a negative tuberculin test = 18/(18+25) = 0.42 = **Negative Predictive Value (NPV)**
- Positive False Ratio: PFR = Pr(B|nonA)
- Negative False Ratio: RFN = Pr(nonA|B)

Independent Events: Conditional Probabilities

- Events A and B are independent if the probability of event B is the same whether or not A has occurred.
- Two events A and B are Independent IF

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$$

- If (and only if) A and B are independent, then:
  - $\Pr(B|A) = \Pr(B|\text{nonA}) = \Pr(B)$
  - $\Pr(A|B) = \Pr(A|\text{nonB}) = \Pr(A)$
- It expressed the independence of the two events: the probability of event B (respectively A) did not depend by the realization of event A (respectively B)
  - Example: if a coin is toss twice the probability to obtain “head” to the second toss is always 0.5 and is not depending if at the first toss we obtained “head” or “tail”.

RECALL!

- **Addition rules:**
  - $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$
  - Pr(A or B) = Pr(A) + Pr(B) – Pr(A and B)
  - Mutually exclusive events:
    - $Pr(A \cup B) = Pr(A) + Pr(B)$
    - Pr(A or B) = Pr(A) + Pr(B):
- **Multiplication Rule:**
  - $Pr(A \cap B) = Pr(A) \cdot Pr(B|A)$
  - Pr(A and B) = Pr(A)·Pr(B|A)
  - Independent events:
    - $Pr(A \cap B) = Pr(A) \cdot Pr(B)$

Random Experiment and Random Variables

RANDOM VARIABLE

- A random variable
  - is a quantification of a probability model that allows to model random data
  - Is a quantity that may take any value of a given range that cannot be predicted exactly but can be described in terms of their probability
    - Discrete: generally assessed by counting
    - Continuous: generally assessed by measurements

RANDOM VARIABLE

- When throwing a die with 6 faces, let X be the random variable defined by:

X = the square of the scores shown on the die

What is the expectation of X?

Solution:

- S = {1, 2^2, 3^2, 4^2, 5^2, 6^2} = {1, 4, 9, 16, 25, 36}
- Each face has a probability of 1/6 of occurring, so:

E(X) = 1*1/6+4*1/6+6*1/6+16*1/6+25*1/6+36*1/6 = 91*1/6

Random Variables

- Outcome of an experiment → a quantitative or a qualitative data
- A random variable is a function that assign a unique numerical value to each outcome of an experiment along with an associated probability.
  - Discrete random variable corresponds to an experiment that has a finite or countable number of outcomes.
  - A continuous random variable is a random variable for which the set of possible outcomes is continuous

RANDOM VARIABLES

- Discrete random variable: take a finite discrete value
  - Infinite number of values
    - Number of hospitalization stay: X = {0, 1, 2, …, n, …}
    - Number of bacteria: X = {0, 1, 2, …, n, …}
  - Finite number of values
    - Number of A blood group in a sample of subjects: X = {0, 1, 2, …, n}
- Continuous random variable: random variable where the data can take infinitely many values
  - Infinite number of values

Discrete Probability Distributions

- Probability Distribution: symbols

$$X : \begin{pmatrix} X_1 & X_2 & \dots & X_n \\ Pr(X_1) & Pr(X_2) & \dots & Pr(X_n) \end{pmatrix}$$

- **Property:** the probabilities that appear in distribution of a finite random variable verify the formula:

$$\sum_{i=1}^{n} Pr(X_i) = 1$$

Discrete Probability Distributions

- The **mean** of discrete probability distribution (called also expected value) is give by the formula:

$$M(X) = \sum_{i=1}^{n} X_i \cdot Pr(X_i)$$

- Represents the weighted average of possible values, each value being weighted by its probability of occurrence.

- Variance: is a weighted average of the squared deviations in X

$$V(X) = \sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)$$

- Standard deviation:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)}$$

DISCRETE PROBABILITY DISTRIBUTIONS

Example: Let X be a random variable represented by the number of otitis episodes in the first two years of life in a community. This random variable has the following distribution:

$$X:\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0.129 & 0.264 & 0.271 & 0.185 & 0.095 & 0.039 & 0.017 \end{pmatrix}$$

What is the expected number (average) of episodes of otitis during the first two years of life?

Discrete Probability Distributions

- What is the **expected number (average)** of episodes of otitis during the first two years of life?

$$X: \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0.129 & 0.264 & 0.271 & 0.185 & 0.095 & 0.039 & 0.017 \end{pmatrix}$$

- $M(X) = 0 \cdot 0.129 + 1 \cdot 0.264 + 2 \cdot 0.271 + 3 \cdot 0.185 + 4 \cdot 0.095 + 5 \cdot 0.039 + 6 \cdot 0.017$
- $M(X) = 0 + 0.264 + 0.542 + 0.555 + 0.38 + 0.195 + 0.102$
- $M(X) = 2.038$

DISCRETE PROBABILITY DISTRIBUTIONS

X_i	$Pr(X_i)$	$X_i*Pr(X_i)$	$X_i-M(X)$	$(X_i-M(X))^2$	$(X_i-M(X))^2*Pr(X_i)$
0	0.129	0	-2.038	4.153	0.536
1	0.264	0.264	-1.038	1.077	0.284
2	0.271	0.542	-0.038	0.001	0.000
3	0.185	0.555	0.962	0.925	0.171
4	0.095	0.38	1.962	3.849	0.366
5	0.039	0.195	2.962	8.773	0.342
6	0.017	0.102	3.962	15.697	0.267
		M(X)=**2.038**			V(X)=**1.967**
					σ(X)=**1.402**

Discrete Probability Distributions By Examples

- **Bernoulli:** head versus tail (two possible outcomes)
- **Binomial:** number of 'head' obtained by throwing a coin of n times
- **Poisson:** number of patients consulted in a emergency office in one day

BERNOULLI DISTRIBUTION

- A random variable whose response are dichotomous is called a Bernoulli random variable
- The experiment has just two possible outcomes (success and failure – dichotomial variable):
  - Gender: boy or girl
  - Results of a test: positive or negative
- Probability of success = p
- Probability of failure = 1-p

X	1	0
Pr(X=x)	p	1-p

- Mean of X: $M(X) = 1 \cdot p + 0 \cdot (1-p)$
- Variance of X: $V(X) = p \cdot (1-p)$

Binomial Distribution

- An experiment is given by repeating a test of *n* times (*n* = known natural number).
- The possible outcomes of each attempt are two events called success and failure
- Let noted with *p* the probability of success and with *q* the probability of failure ($q = 1 - p$)
- The *n* repeated tests are independent
- In a binomial experiment:
  - It consists of a fixed number n of identical experiments
  - There are only two possible outcomes in each experiments, denoted by S (success) and F (failure)
  - The experiments are independent with the same probability of S (denoted p)
- ${}_4C_2$ = 4 choose 2 (combination choosing 2 from 4)

BINOMIAL DISTRIBUTION

Mean	$M(X) = n \cdot p$
Variance	$V(X) = n \cdot p \cdot q$
Standard deviation	$\sigma(X) = \sqrt{(n \cdot p \cdot q)}$

- The number of successes X obtained by performing the test n times is a random variable of n and p parameters and is noted as Bi(n,p)
- The random variable X can take the following values: 0, 1, 2,...n
- Probability that X to be equal with a value k is given by the formula: $$\Pr(X = k) = C_n^k p^k q^{n-k}$$

  where: $$C_n^k = \frac{n!}{k! \cdot (n-k)!}$$

Binomial Distribution

- Suppose that 90% of adults with joint pain report symptomatic relief with a specific medication. If the medication is given to 10 new patients with joint pain, what is the probability that it is effective in exactly 7 subjects?
- Assumptions of the binomial distribution model?
  - The outcome is pain relief (yes or no), and we will consider that the *pain relief* is a success
  - The probability of success for each subject is 0.9 (p=0.9)
  - The final assumption is that the subjects are independent, and it is reasonable to assume that this is true.
- $Pr(X=7) = {}_{10}C_7 \cdot 0.9^7 \cdot (1-0.9)^{10-7} = 0.0574$ → there is a 5.74% chance that exactly 7 of 10 patients will report pain relief when the probability that any one reports relief is 90%.

BINOMIAL DISTRIBUTION

$$\Pr(X = k) = C_n^k p^k q^{n-k}$$

- What is the probability to have 2 boys in a family with 5 children if the probability to have a baby-boy for each birth is equal to 0.47 and if the sex of children successive born in the family is considered an independent random variable?

- p=0.47
- q=1-0.47=0.53
- n=5
- k=2
- **Pr(X=2)=10·0.47^2·0.53^3**
- **Pr(X=2) = 0.33**

$$C_5^2 = \frac{5!}{2!\cdot(5-2)!} = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{2\cdot 1\cdot (3\cdot 2\cdot 1)} = \frac{120}{12} = 10$$

POISSON DISTRIBUTION

- Random Poisson variable take a countable infinity of values (0,1,2,...,k,...) that is the number of achievements of an event within a given range of time or place
  - number of entries per year in a given hospital
  - white blood cells on smear
  - number of decays of a radioactive substance in a given time T

Poisson Distribution

- POISSON random variable:
  - Is characterized by theoretical parameter θ (expected average number of achievement for a given event in a given range)
- Symbol: Po(θ)
- Poisson Distribution:

$$X:\begin{pmatrix} k \\ e^{-\theta}\cdot\dfrac{\theta^k}{k!} \end{pmatrix}$$

$$\Pr(X = k) = \frac{e^{-\theta}\cdot\theta^k}{k!}$$

- Mean of expected values: M(X) = θ
- Variance: V(X) = θ

Poisson Distribution

- The mortality rate for specific viral pathology is 7 per 1000 cases. What is the probability that in a group of 400 people this pathology to cause 5 deaths?
- $n=400$
- $p=7/1000=0.007$
- $\theta=n\cdot p=400\cdot 0.007=2.8$
- $e=2.718281828=2.72$

$$\mathbf{Pr(X=5) = (2.72^{-2.8}\cdot 2.8^{5})/(5\cdot 4\cdot 3\cdot 2\cdot 1) = 10.45/120}$$

$$\mathbf{Pr(X=5) = 0.09}$$

Recall

- Random variables could be discrete or continuous.
- For random variables we have:
  - Discrete probability distributions
  - Continuous probability distribution
- It is possible to compute the expected values (means), variations and standard deviation for discrete and random variables.

Thank you!