Probablity

Probablity Vs Statistics

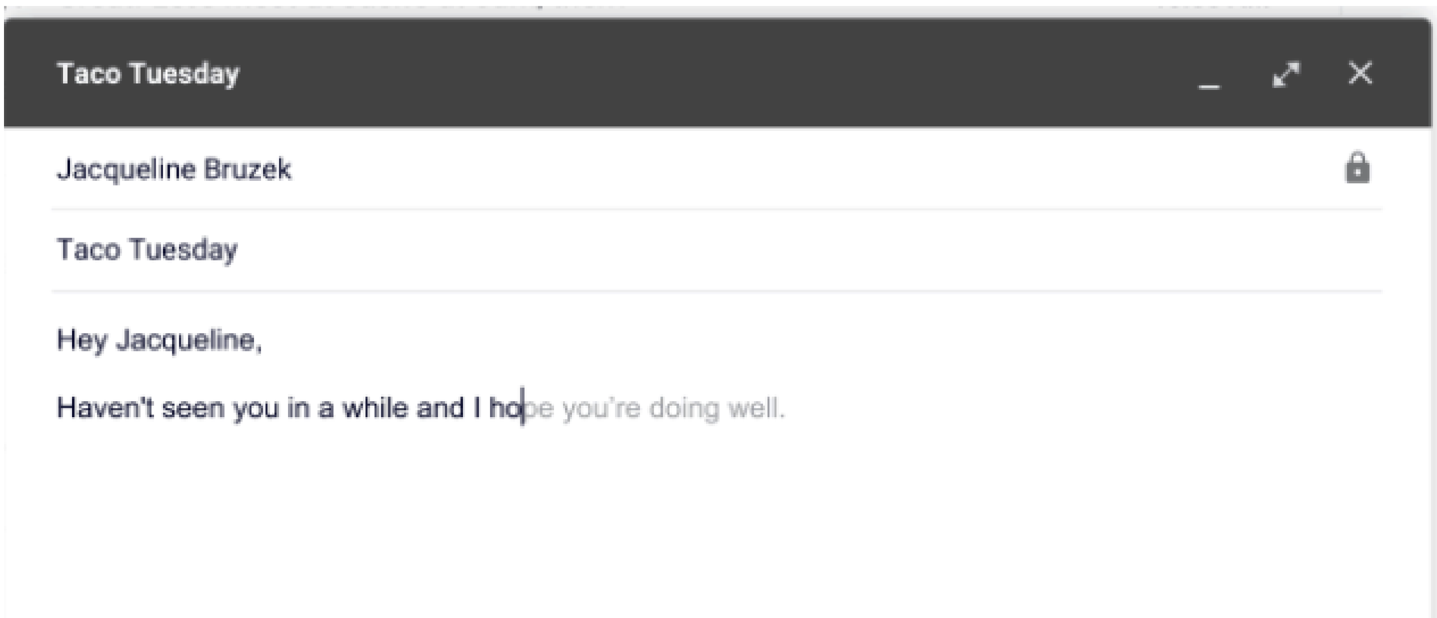
Probablity Vs Statistics

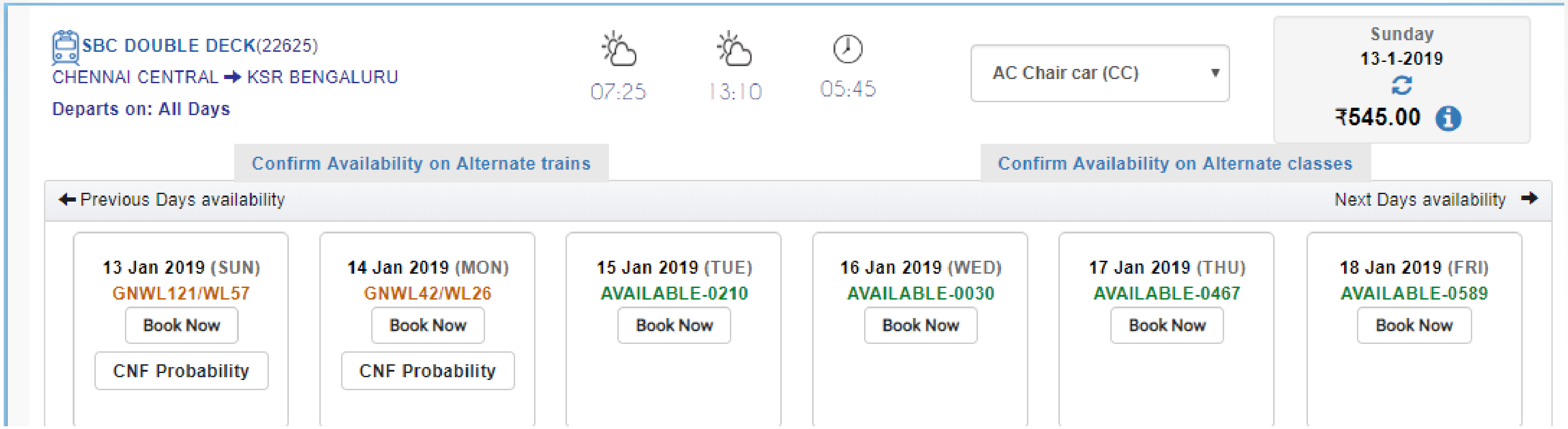
- Probability – Predict the likelihood of a future event
- Statistics – Analyze the past events

- Probability – What will happen in a given ideal world?
- Statistics – How ideal is the world?

Probablity Vs Statistics

Probability is the basis of inferential statistics.


Taco Tuesday
Jacqueline Bruzek
Taco Tuesday
Hey Jacqueline,
Haven't seen you in a while and I hope you're doing well.


SBC DOUBLE DECK(22625)
CHENNAI CENTRAL → KSR BENGALURU
Departs on: All Days
07:25
13:10
05:45
AC Chair car (CC)
Sunday
13-1-2019
₹545.00
Confirm Availability on Alternate trains
Confirm Availability on Alternate classes
Previous Days availability
Next Days availability
13 Jan 2019 (SUN)
GNWL121/WL57
Book Now
CNF Probability
14 Jan 2019 (MON)
GNWL42/WL26
Book Now
CNF Probability
15 Jan 2019 (TUE)
AVAILABLE-0210
Book Now
16 Jan 2019 (WED)
AVAILABLE-0030
Book Now
17 Jan 2019 (THU)
AVAILABLE-0467
Book Now
18 Jan 2019 (FRI)
AVAILABLE-0589
Book Now

SBC DOUBLE DECK(22625)
CHENNAI CENTRAL ➜ KSR BENGALURU
Departs on: All Days

07:25 13:10 05:45

AC Chair car (CC)

Sunday
13-1-2019
₹545.00

Confirm Availability on Alternate trains

Confirm Availability on Alternate classes

← Previous Days availability

Next Days availability →

13 Jan 2019 (SUN)	14 Jan 2019 (MON)	15 Jan 2019 (TUE)	16 Jan 2019 (WED)	17 Jan 2019 (THU)	18 Jan 2019 (FRI)
GNWL121/WL57	GNWL42/WL26	AVAILABLE-0210	AVAILABLE-0030	AVAILABLE-0467	AVAILABLE-0589
Book Now	Book Now	Book Now	Book Now	Book Now	Book Now
CNF Probability	CNF Probability				

SBC DOUBLE DECK(22625)

CHENNAI CENTRAL (MAS) ➜ KSR BENGALURU (SBC)
AC Chair car (CC)

Current availability as on 13-1-2019: GNWL121/WL57

Probability of Confirmation: 65% *

Last Year Confirmation Trends for Same Period

Journey Date	Booking Confirmed Upto	Journey Date	Booking Confirmed Upto
06 Jan 2018	WL/34	07 Jan 2018	WL/19
08 Jan 2018	WL/47	12 Jan 2018	WL/49
13 Jan 2018	WL/97	14 Jan 2018	WL/21
15 Jan 2018	WL/45	16 Jan 2018	WL/33
17 Jan 2018	WL/27	18 Jan 2018	WL/8
20 Jan 2018	WL/65		

Click Here to Check CNF Availability on alternate Train

Probability - Applications

8 National Vital Statistics Reports, Vol. 54, No. 14, April 19, 2006

Table 1. Life table for the total population: United States, 2003

Age	Probability of dying between ages x to $x+1$	Number surviving to age x	Number dying between ages x to $x+1$
	$q(x)$	$l(x)$	$d(x)$
0–1	0.006865	100,000	687
1–2	0.000469	99,313	47
2–3	0.000337	99,267	33
3–4	0.000254	99,233	25
4–5	0.000194	99,208	19
5–6	0.000177	99,189	18
6–7	0.000160	99,171	16

Insurance industry uses probabilities in actuarial tables for setting premiums and coverages.

Assigning Probabilities

Classical Method – *A priori* or Theoretical

Probability can be determined prior to conducting any experiment.

$$P(E) = \frac{\#\ of\ \ outcomes\ in\ which\ the\ event\ occurs}{total\ possible\ \#\ of\ \ outcomes}$$

Example: Tossing of a fair die

Assigning Probabilities

Empirical Method – *A posteriori* or Frequentist

Probability can be determined post conducting a thought experiment.

$$P(E) = \frac{\#\ of\ times\ an\ event\ occurred}{total\ \#\ of\ opportunities\ for\ the\ event\ to\ have\ occurred}$$

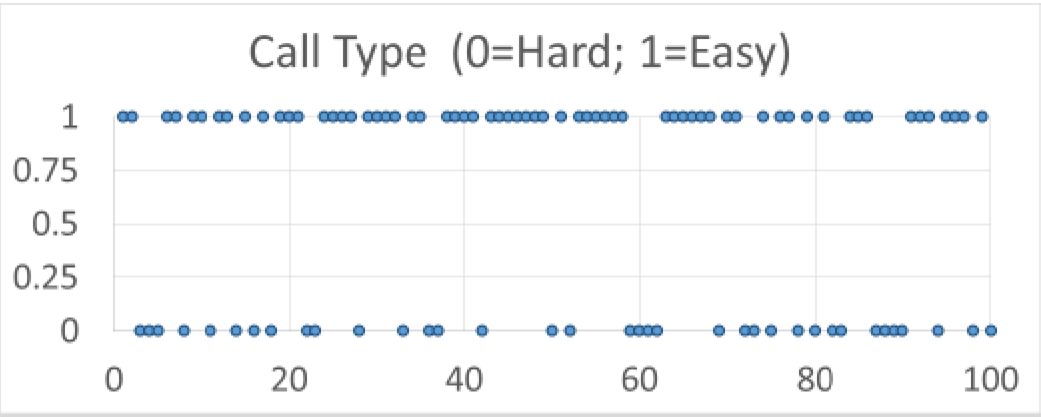
Example: Tossing of a weighted die…well!, even a fair die. The larger the number of experiments, the better the approximation.

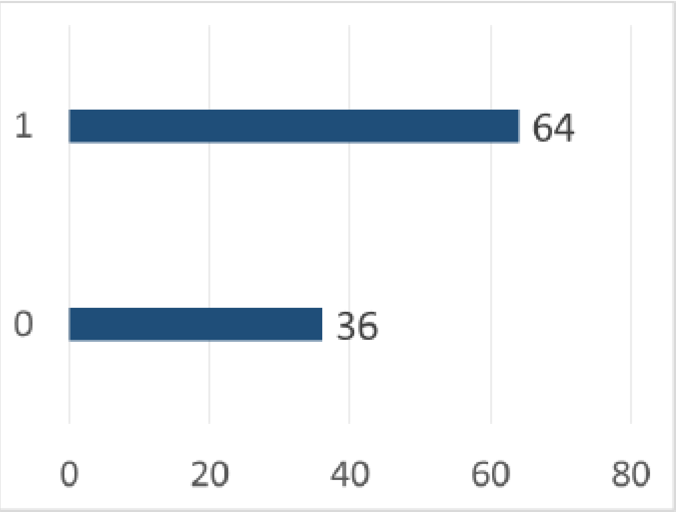
This is the most used method in statistical inference.

Assigning Probabilities

Empirical Method – *A posteriori* or Frequentist

100 calls handled by an agent at a call centre

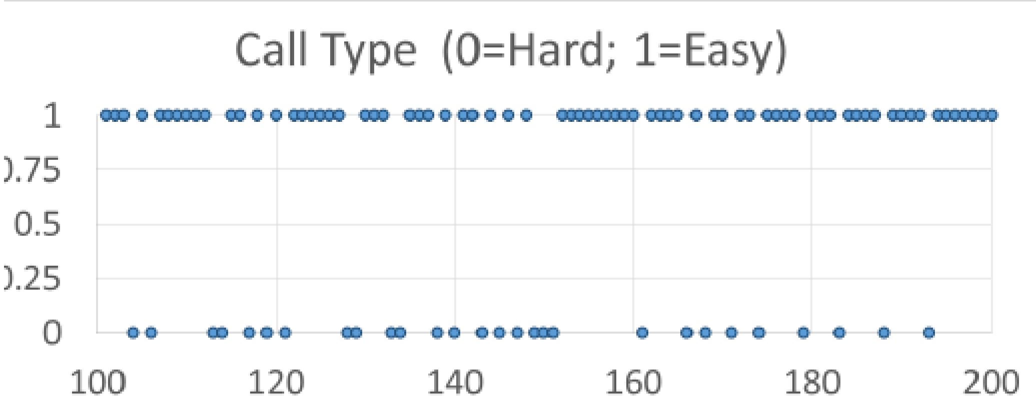


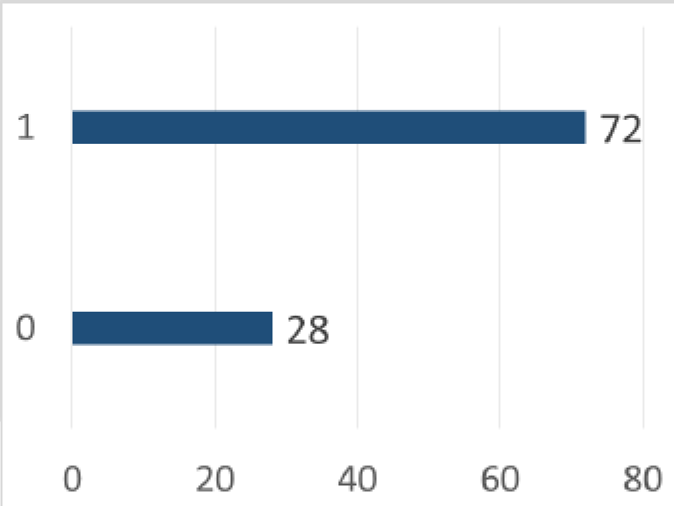


Assigning Probabilities

Empirical Method – *A posteriori* or Frequentist

Next 100 calls handled by an agent at a call centre

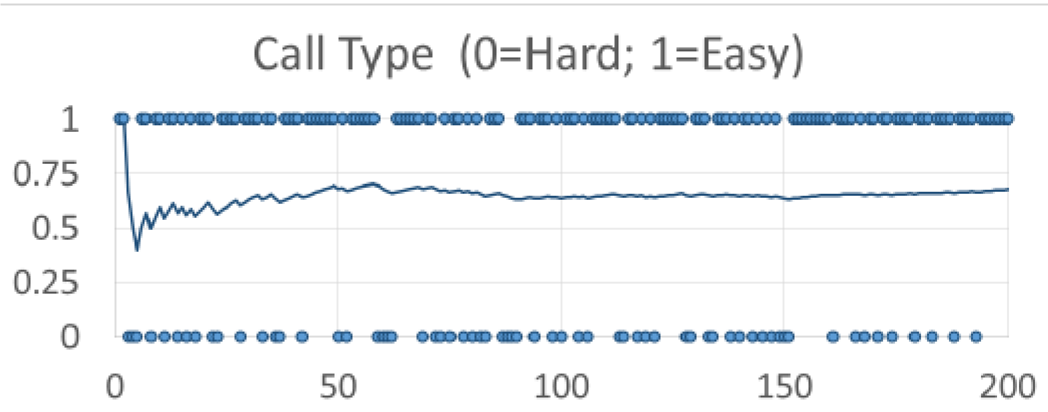


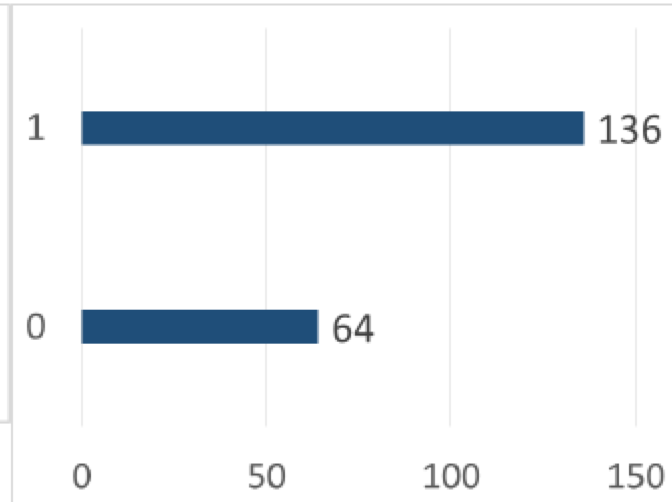


Assigning Probabilities

Empirical Method – *A posteriori* or Frequentist

Averages over the long run





P(easy) = 0.7

Assigning Probabilities

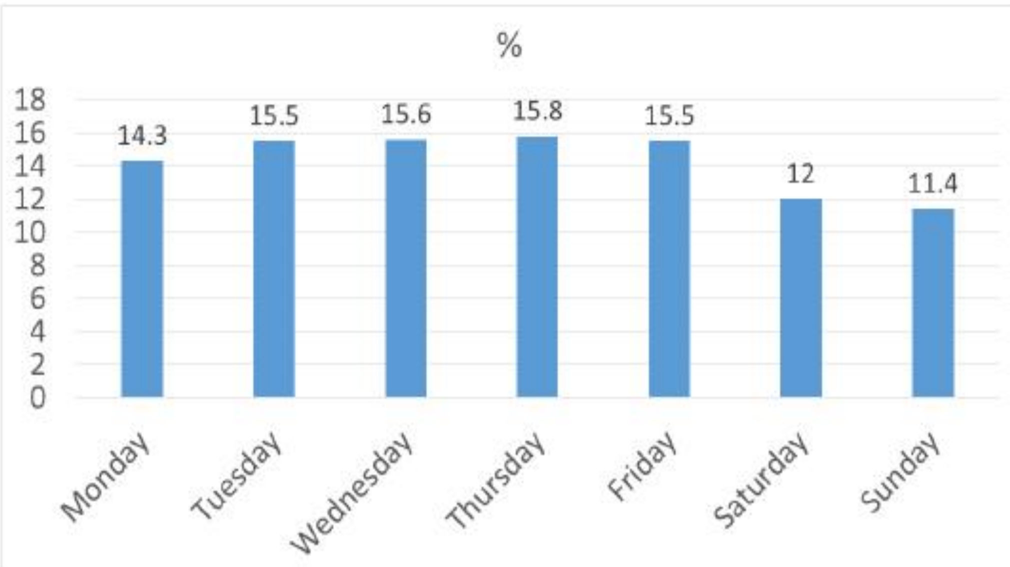
Subjective Method

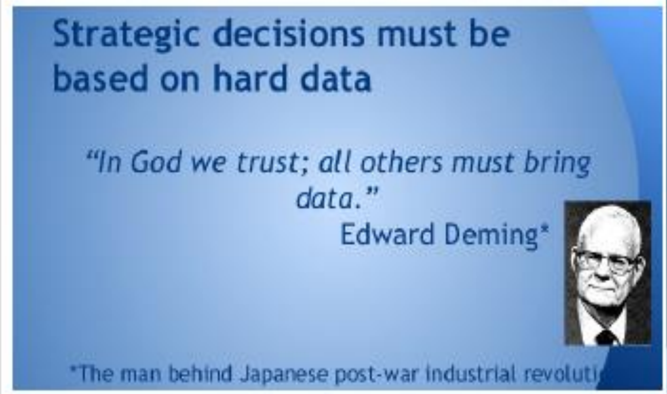
Based on feelings, insights, knowledge, etc. of a person.

What is the probability of rain tomorrow?

Assigning Probabilities

What is the probability of a baby being born on a Sunday?



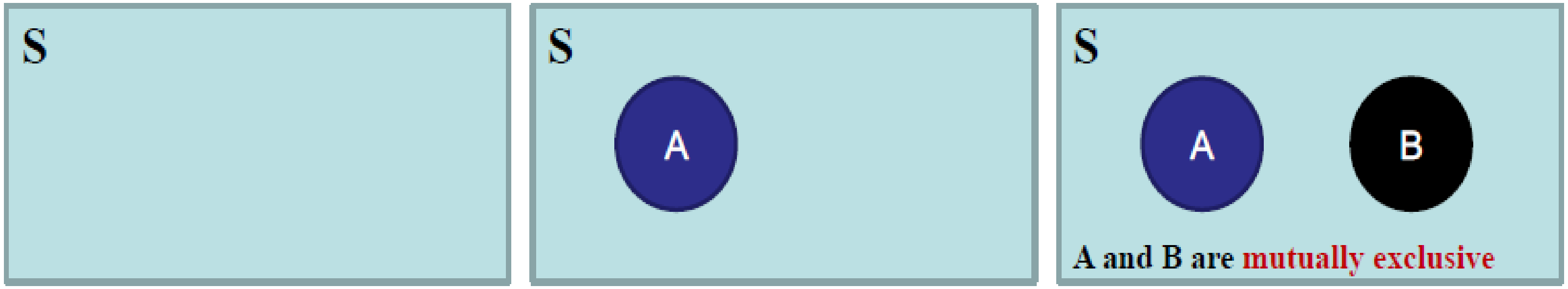


Probability - Terminology

Sample Space – Set of all possible outcomes, denoted S.

Event – A subset of the sample space.

Probability - Rules



$$P(S) = 1$$

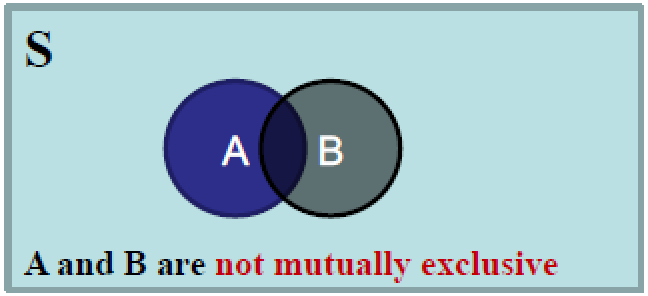
$$0 \leq P(A) \leq 1$$

$$P(A \text{ or } B) = P(A) + P(B)$$

Area of the rectangle denotes sample space, and since probability is associated with area, it cannot be negative.

Mutually Exclusive – If event A happens, event B cannot.

Probability - Rules



$$P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$$

Example

Event A – Customers who default on loans

Event B – Customers who are High Net Worth Individuals

Probability - Rules

Independent Events – Outcome of event B is not dependent on the outcome of event A.

Probability of customer B defaulting on the loan is not dependent on default (or otherwise) by customer A.

$$P(A \text{ and } B) = P(A) * P(B)$$

If the probability of getting an *easy* call is 0.7, what is the probability that the next 3 calls will be *easy*?

$$P(easy_1 \text{ and } easy_2 \text{ and } easy_3) = 0.7^3 = 0.343$$

Probability - Question

A basketball team is down by 2 points with only a few seconds remaining in the game. Given that:

- Chance of making a 2-point shot to tie the game = 50%
- Chance of winning in overtime = 50%
- Chance of making a 3-point shot to win the game = 30%

What should the coach do: go for 2-point or 3-point shot?

What are the assumptions, if any?

Probability - Types

Contingency table summarizing 2 variables, *Loan Default* and *Age*:

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	10,503	27,368	259	38,130
	Yes	3,586	4,851	120	8,557
	Total	14,089	32,219	379	46,687

Probability - Types

Convert it into probabilities:

		Age			
		Young	**Middle-aged**	**Old**	**Total**
Loan Default	**No**	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

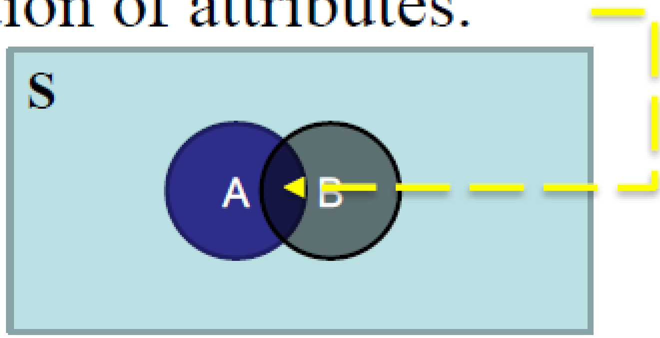
Probability - Types

Joint Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

Probability describing a combination of attributes.

P(Yes **and** Young) = 0.077

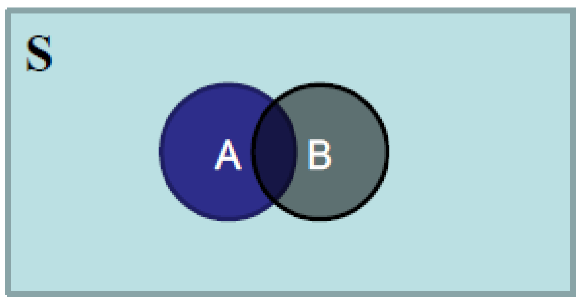


Probability - Types

Union Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

P(Yes **or** Young) = P(Yes) + P (Young) – P(Yes and Young) = 0.184 + 0.302 – 0.077 = 0.409

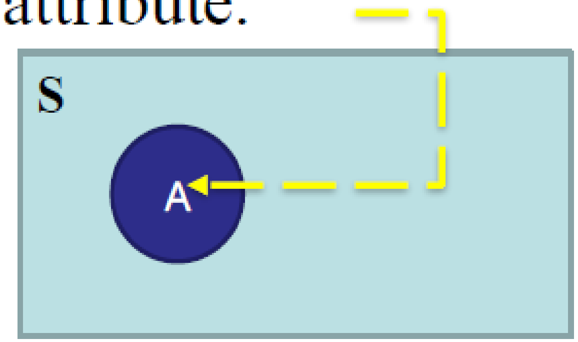


Probability - Types

Marginal Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

Probability describing a single attribute.



P(No) = 0.816

P(Old) = 0.008

Probability - Types

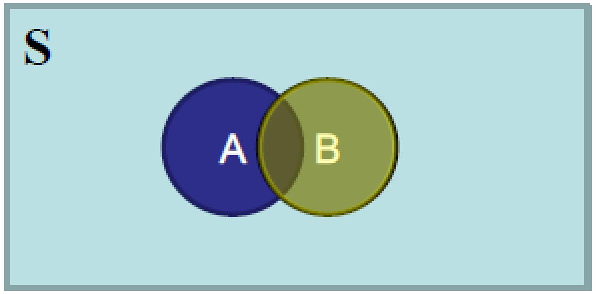
Conditional Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

Probability of A occurring **given that** B has occurred.

The sample space is restricted to a single row or column.

This makes rest of the sample space irrelevant.



Probability - Types

Conditional Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

What is the probability that a person will not default on the loan payment **given** she is middle-aged?

P(No | Middle-Aged) = 0.586/0.690 = 0.85

Note that this is the ratio of

Probability - Types

Conditional Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

What is the probability that a person will not default on the loan payment **given** she is middle-aged?

P(No | Middle-Aged) = 0.586/0.690 = 0.85

Note that this is the ratio of **Joint Probability** to **Marginal Probability**, i.e., $P(A|B) = \frac{P(A \; and \; B)}{P(B)}$

Probability - Types

Conditional Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

What is the probability that a person will not default on the loan payment **given** she is middle-aged?

P(No | Middle-Aged) = 0.586/0.690 = 0.85

Note that this is the ratio of **Joint Probability** to **Marginal Probability**, i.e., $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$

P(Middle-Aged | No) =

Probability - Types

Conditional Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

What is the probability that a person will not default on the loan payment **given** she is middle-aged?

P(No | Middle-Aged) = 0.586/0.690 = 0.85

Note that this is the ratio of **Joint Probability** to **Marginal Probability**, i.e., $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$

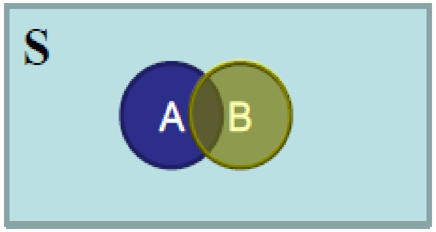
P(Middle-Aged | No) = 0.586/0.816 = 0.72 (Order Matters)

Probability - Types

Conditional Probability – Visualizing using Probability Tables and Venn Diagrams

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	10,503	27,368	259	38,130
	Yes	3,586	4,851	120	8,557
	Total	14,089	32,219	379	46,687

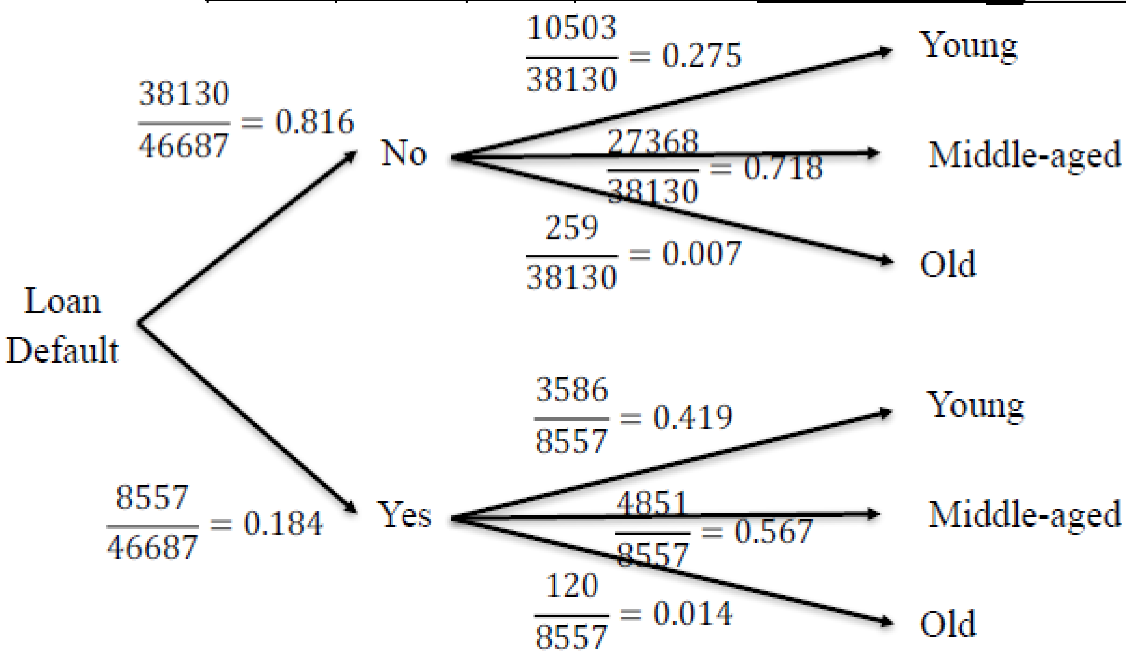
		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000



Probability - Types

Conditional Probability – Visualizing using Probability Trees

		Age (Numbers)				Age (Probabilities)			
		Young	Middle-aged	Old	Total	Young	Middle-aged	Old	Total
Loan Default	No	10,503	27,368	259	38,130	0.225	0.586	0.005	0.816
	Yes	3,586	4,851	120	8,557	0.077	0.104	0.003	0.184
	Total	14,089	32,219	379	46,687	0.302	0.690	0.008	1.000



Find

- P(Old and Yes)
- P(Yes and Old)
- P(Old)
- P(Yes)
- P(Old | Yes)
- P(Yes | Old)
- P(Young | No)

Probability - Types

Attention Check

Identify the type of probability in each of the below cases:

1. P(Old and Yes)
2. P(Yes and Old)
3. P(Old)
4. P(Yes)
5. P(Old | Yes)
6. P(Yes | Old)
7. P(Young | No)
8. P(Middle-aged or No)
9. P(Old or Young)

		Age (Probabilities)			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	**0.816**
	Yes	0.077	0.104	0.003	**0.184**
	Total	**0.302**	**0.690**	**0.008**	**1.000**

Probability - Types

Attention Check

Identify the type of probability in each of the below cases:

1. P(Old and Yes)
2. P(Yes and Old)
3. P(Old)
4. P(Yes)
5. P(Old | Yes)
6. P(Yes | Old)
7. P(Young | No)
8. P(Middle-aged or No)
9. P(Old or Young)

		Age (Probabilities)			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	**0.816**
	Yes	0.077	0.104	0.003	**0.184**
	Total	**0.302**	**0.690**	**0.008**	**1.000**

1 and 2: **Joint**; 3 and 4: **Marginal**; 5, 6 and 7: **Conditional**; 8 and 9: **Union**

Probability - Types

Conditional Probability

$$P(A|B) = \frac{P(A \text{ } and \text{ } B)}{P(B)} \Rightarrow P(A \text{ } and \text{ } B) = P(B) * P(A|B)$$

Similarly

$$P(B|A) = \frac{P(A \text{ } and \text{ } B)}{P(A)} \Rightarrow P(A \text{ } and \text{ } B) = P(A) * P(B|A)$$

Equating, we get

$$P(A|B) * P(B) = P(A) * P(B|A)$$

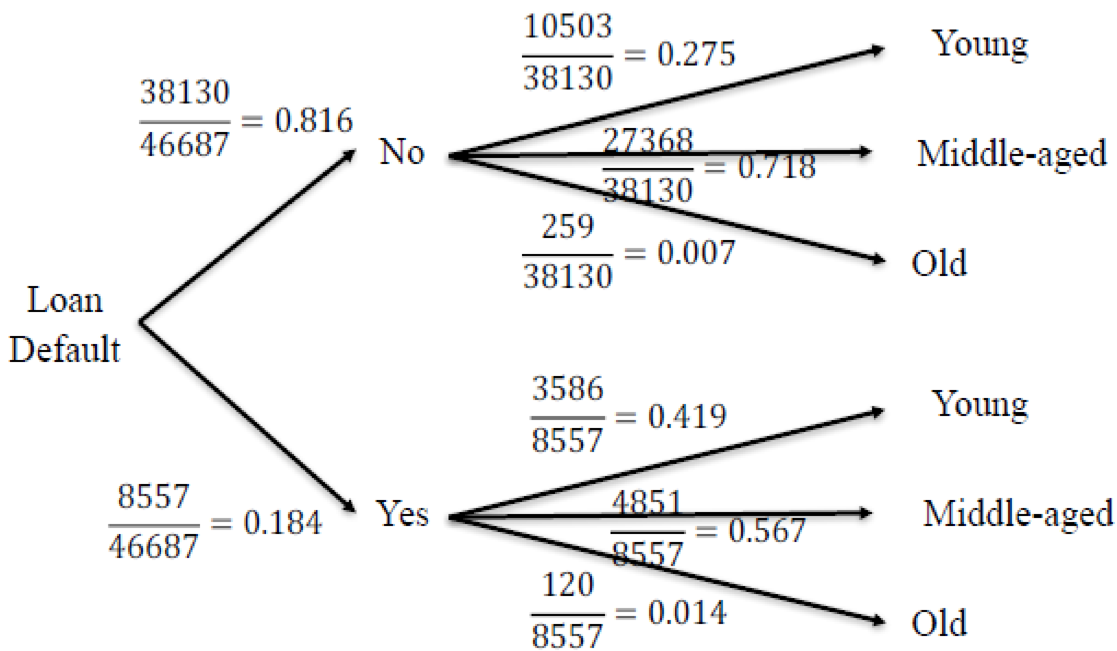
$$\therefore P(A|B) = \frac{P(A) * P(B|A)}{P(B)}$$

Probability - Types

Conditional Probability – Visualizing using Probability Trees

		Age (Probabilities)			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

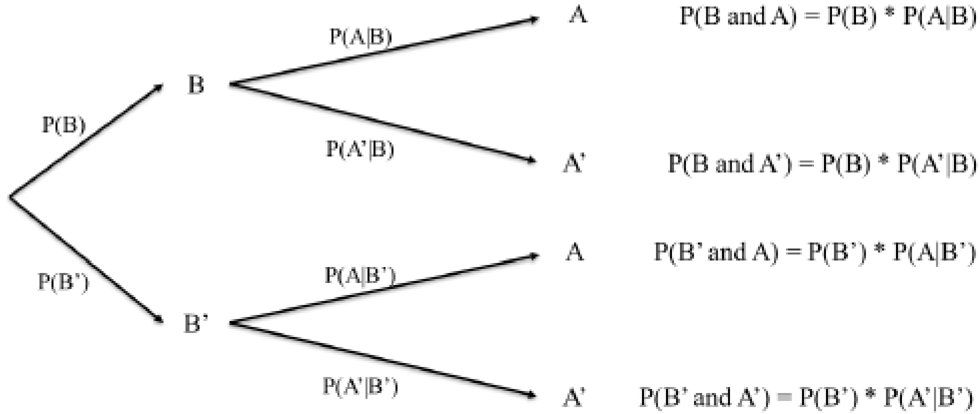
$$P(A|B) = \frac{P(A) * P(B|A)}{P(B)}$$



Now find P(Yes | Old)

Probability - Types

Conditional Probability -> Bayes’ Theorem



$$P(B|A) = \frac{P(B) * P(A|B)}{P(A)} = \frac{P(A|B) * P(B)}{P(A|B) * P(B) + P(A|not\ B) * P(not\ B)}$$

Note B’ means “not B”

Bayes' Theorem

Bayes' Theorem allows you to find reverse probabilities, and to allow **revision of original probabilities** with new information.

Case – Clinical trials

Epidemiologists claim that probability of breast cancer among Caucasian women in their mid-50s is 0.005. An established test identified people who had breast cancer and those that were healthy. A new mammography test in clinical trials has a probability of 0.85 for detecting cancer correctly. In women without breast cancer, it has a chance of 0.925 for a negative result. If a 55-year-old Caucasian woman tests positive for breast cancer, what is the probability that she in fact has breast cancer?

Bayes' Theorem

Case – Clinical trials

P(Cancer) = 0.005

P(Test positive | Cancer) = 0.85 (*aka* Prior Probability)

P(Test negative | No cancer) = 0.925

P(Cancer | Test positive) = ? (*aka* Posterior or Revised Probability)

$$P(Cancer|Test+) = \frac{P(Cancer) * P(Test+|Cancer)}{P(Test+|Cancer) * P(Cancer) + P(Test+|No\ cancer) * P(No\ cancer)}$$

$$= \frac{0.005 * 0.85}{0.85 * 0.005 + 0.075 * 0.995} = \frac{0.00425}{0.078875} = 0.054$$

Homework

Draw a Probability Table and a Probability Tree for the above case.

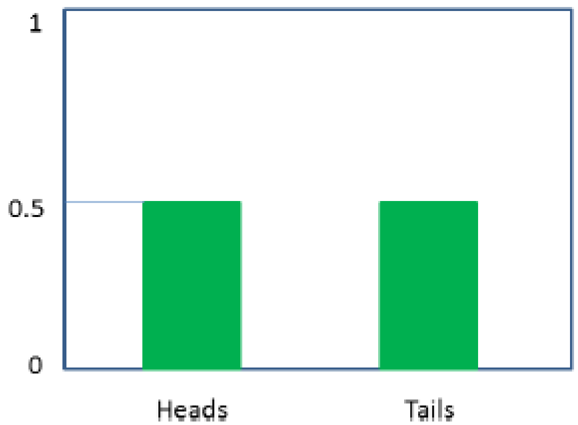
Analyzing attributes

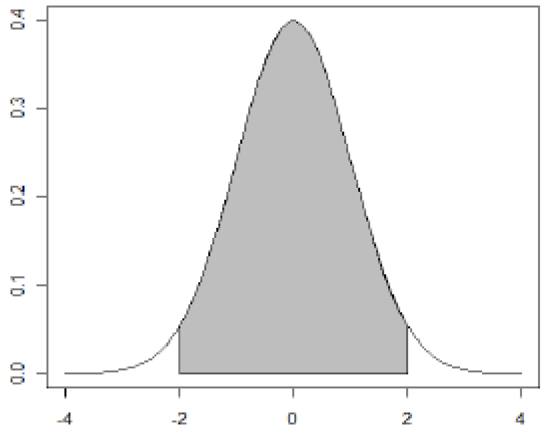
PROBABILITY DISTRIBUTIONS

Random variable

- A variable that can take multiple values with different probabilities.
- The mathematical function describing these possible values along with their associated probabilities is called a probability distribution.

Discrete and Continuous





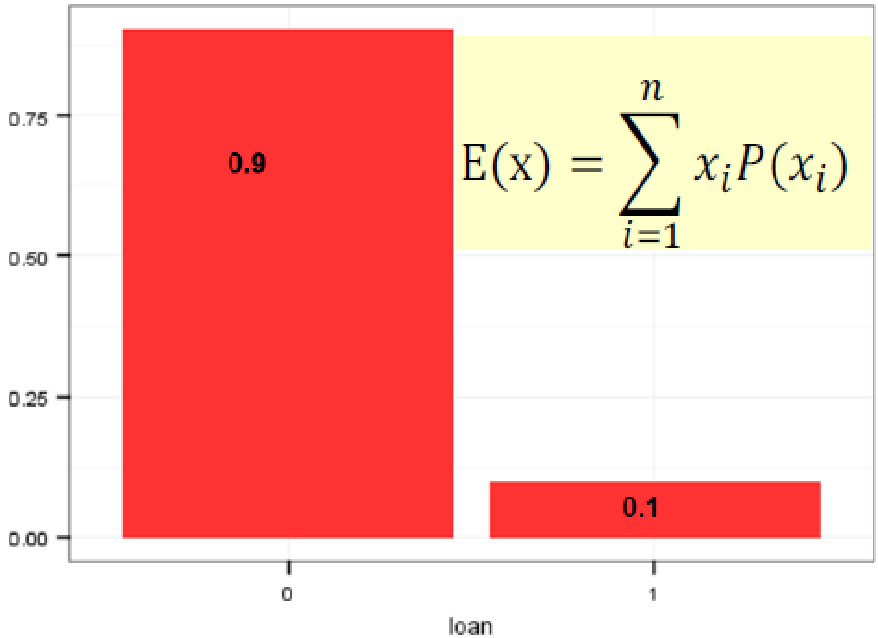
Countable

Measurable

Can any function be a probability distribution?

Discrete Distributions	Continuous Distributions
Probability that X can take a specific value x is $P(X = x) = p(x)$.	Probability that X is between two points a and b is $P(a \leq X \leq b) = \int_a^b f(x)dx$.
It is non-negative for all real x.	It is non-negative for all real x.
The sum of $p(x)$ over all possible values of x is 1, i.e., $\sum p(x) = 1$.	$\int_{-\infty}^{\infty} f(x)dx = 1$
Probability Mass Function	**Probability Density Function**

Expectation: Discrete



$$E(x) = \sum_{i=1}^{n} x_i P(x_i)$$

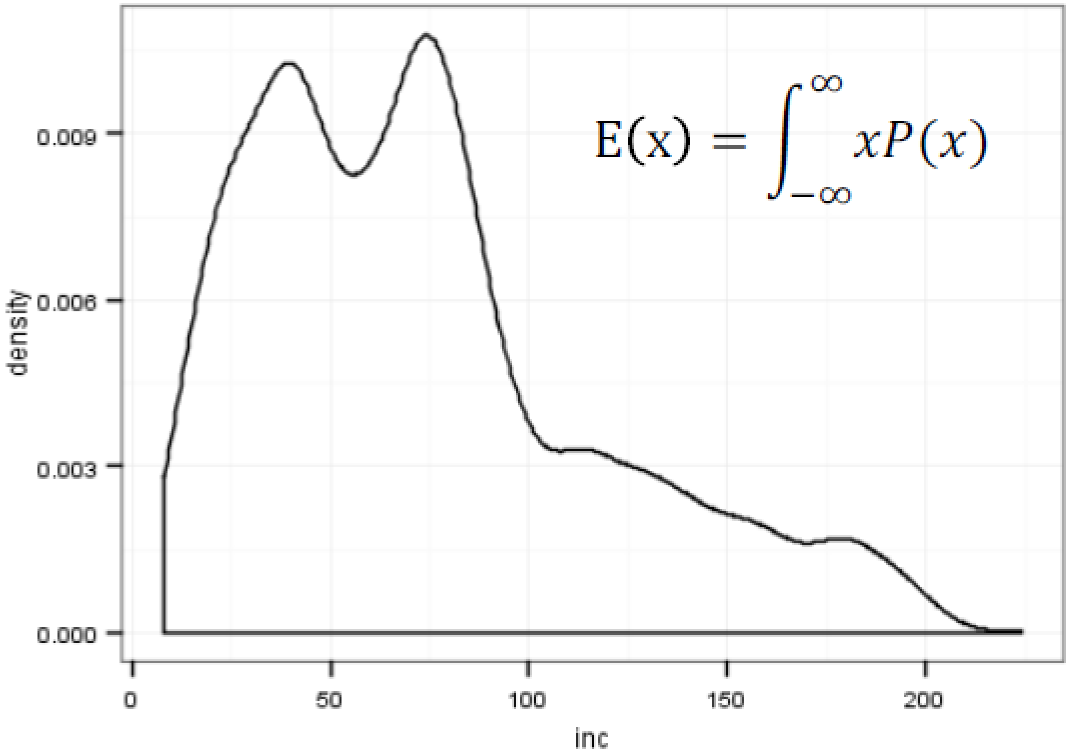
Recall anything like this?

Salary (BHD)	100	345	1000	9833
Frequency, f	10	1	10	2
Probability	0.43	0.04	0.43	0.09

$$\text{Mean, } \mu = \frac{\Sigma x}{n} = \frac{\Sigma f x}{\Sigma f} = \frac{100X10+345X1+1000X10+9833X2}{10+1+10+2} = 1348$$

$$\text{Expectation, } E(X) = 100 * 0.43 + 345 * 0.04 + 1000 * 0.43 + 9833 * 0.09 = 1348$$

Expectation: Continuous



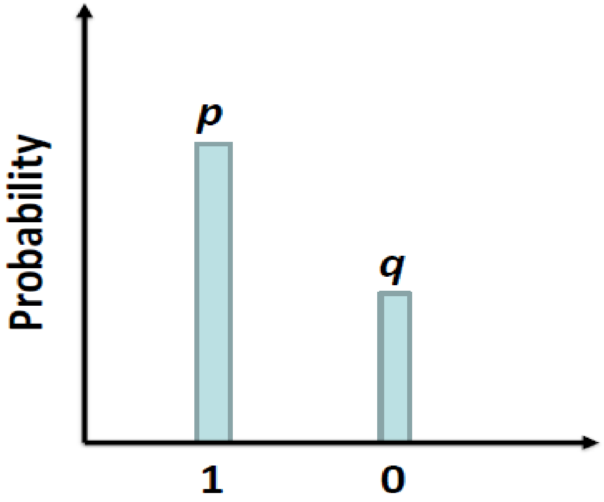
$$E(x) = \int_{-\infty}^{\infty} x P(x)$$

SOME COMMON DISTRIBUTIONS

Bernoulli

There are two possibilities (loan taker or non-taker) with probability *p* of success and *1-p* of failure

- Expectation: p
- Variance: $p(1-p)$ or pq, where $q=1-p$



Geometric Distribution

Number of independent and identical Bernoulli trials needed to get ONE success, e.g., number of people I need to call for the first person to accept the loan.

Geometric Distribution

PMF*, $P(X = r) = q^{r-1}p$ — *(r-1)* failures followed by ONE success.

$P(X > r) = q^r$ — Probability you will need more than *r* trials to get the first success.

CDF**, $P(X \leq r) = 1 - q^r$ — Probability you will need *r* trials or less to get your first success.

$$E(X) = \frac{1}{p} \qquad Var(X) = \frac{q}{p^2}$$

* Probability Mass Function ** Cumulative Distribution Function

Geometric Distribution

- You run a series of independent trials.
- There can be either a success or a failure for each trial, and the probability of success is the same for each trial.
- The main thing you are interested in is how many trials are needed in order to get the first successful outcome.

Binomial Distribution

If there are two possibilities with probability *p* for success and *q* for failure, and if we perform *n* trials, the probability that we see *r* successes is

PMF, $\mathrm{P(X = r)} = C_r^n p^r q^{n-r}$

CDF, $\mathrm{P(X \leq r)} = \sum_{i=0}^{r} C_i^n p^i q^{n-i}$

Binomial Distribution

$$E(X) = np \qquad\qquad Var(X) = npq$$

When to use?

- You run a series of independent trials.
- There can be either a success or a failure for each trial, and the probability of success is the same for each trial.
- There are a finite number of trials, and you are interested in the number of successes or failures.

Poisson Distribution

Probability of getting 15 customers requesting for loans in a given day given on average we see 10 customers

$\lambda = 10 \; and \; r = 15$

PMF, $\mathrm{P(X = r)} = \frac{e^{-\lambda}\lambda^{r}}{r!}$

CDF, $\mathrm{P(X \leq r)} = e^{-\lambda} \sum_{i=0}^{r} \frac{\lambda^{i}}{i!}$

Poisson Distribution

$E(X) = \lambda$ Can be equated to *np* of Binomial if *n* is large (>50) and *p* is small (<0.1)

$Var(X) = \lambda$ Can be equated to *npq* of Binomial in the above situation.

When to use?

- Individual events occur at random and independently in a given interval (time or space).
- You know the mean number of occurrences, λ, in the interval or the rate of occurrences, and it is finite.

Poisson Distribution

The probability that no customer will visit the store in one day

$$P(X=0) = \frac{e^{-\lambda}\lambda^{0}}{0!} = e^{-\lambda}$$

Probability that she will not have a customer for *n* days

$$e^{-n\lambda}$$

Exponential Distribution

Probability that a customer will visit in *n* days: $1 - e^{-n\lambda}$

$$CDF = 1 - e^{-n\lambda}, n \geq 0$$
$$PDF = \lambda e^{-n\lambda}, n \geq 0$$

Distributions

Geometric:	For estimating number of attempts before first success
Binomial:	For estimating number of successes in *n* attempts
Poisson:	For estimating *n* number of events in a given time period when on average we see *m* events
Exponential:	Time between events

Probability Distributions

Here are a few scenarios. Identify the distribution and calculate expectation, variance and the required probabilities.

Q1. A man is bowling. The probability of him knocking all the pins over is 0.3. If he has 10 shots, what is the probability he will knock all the pins over less than 3 times?

Q2. On average, 1 bus stops at a certain point every 15 minutes. What is the probability that no buses will turn up in a single 15 minute interval?

Q3. 20% of cereal packets contain a free toy. What is the probability you will need to open fewer than 4 cereal packets before finding your first toy?

Probability Distributions

Solutions

A man is bowling. The probability of him knocking all the pins over is 0.3. If he has 10 shots, what is the probability he will knock all the pins over less than 3 times?

X ~ B(10,0.3); n=10, p=0.3, q=1-0.3=0.7, r=0, 1, 2 (< 3)

E(X) = np = 3

Var(X) = npq = 2.1

$$P(X = r) = {}^{n}C_{r} p^{r} q^{n-r}$$

P(X=0) = 0.028; P(X=1) = 0.121; P(X=2) = 0.233

∴ P(X<3) = 0.028 + 0.121 + 0.233 = 0.382

- Additional References

What is a Probability Distribution?

A probability distribution is a table or an equation that links each outcome of a **statistical experiment** with its probability of occurrence.

Probability Distribution Prerequisites

To understand probability distributions, it is important to understand variables. random variables, and some notation.

- A **variable** is a symbol (*A*, *B*, *x*, *y*, etc.) that can take on any of a specified set of values.
- When the value of a variable is the outcome of a statistical experiment, that variable is a **random variable**.

What is a Probability Distribution?

- X represents the random variable X.
- P(X) represents the probability of X.
- P(X = x) refers to the probability that the random variable X is equal to a particular value, denoted by x.

Example:

P(X = 1) refers to the probability that the random variable X is equal to 1.

Probability Distributions

An example will make clear the relationship between random variables and probability distributions. Suppose you flip a coin two times. This simple statistical experiment can have four possible outcomes: HH, HT, TH, and TT. Now, let the variable X represent the number of Heads that result from this experiment. The variable X can take on the values 0, 1, or 2. In this example, X is a random variable; because its value is determined by the outcome of a statistical experiment.

A **probability distribution** is a table or an equation that links each outcome of a statistical experiment with its probability of occurrence. Consider the coin flip experiment described above. The table below, which associates each outcome with its probability, is an example of a probability distribution.

Number of heads	Probability
0	0.25
1	0.50
2	0.25

The above table represents the probability distribution of the random variable X.

Cumulative Probability Distributions

A **cumulative probability** refers to the probability that the value of a random variable falls within a specified range.

Let us return to the coin flip experiment. If we flip a coin two times, we might ask: What is the probability that the coin flips would result in one or fewer heads? The answer would be a cumulative probability. It would be the probability that the coin flip experiment results in zero heads plus the probability that the experiment results in one head.

$$P(X \leq 1) = P(X = 0) + P(X = 1) = 0.25 + 0.50 = 0.75$$

Like a probability distribution, a cumulative probability distribution can be represented by a table or an equation. In the table below, the cumulative probability refers to the probability than the random variable X is less than or equal to x.

Number of heads: x	**Probability:** $P(X = x)$	**Cumulative Probability:** $P(X \leq x)$
0	0.25	0.25
1	0.50	0.75
2	0.25	1.00

Applications
- Binomial distribution (practical life)

Explanation:

- Binomial distribution can be used in any task which requires repeating the same experiment more than once and calculating the probability of a specified number of outcomes.

For example we can use it in a question like:

- *Find the probability that in 5 tosses of a fair dice at least 2 numbers will be prime.*

Binomial Probabilities Statistical Tables

n = 10

x	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
0	0.599	0.349	0.197	0.107	0.056	0.028	0.013	0.006	0.003	0.001	0	0	0	0	0	0	0	0	0
1	0.914	0.736	0.544	0.376	0.244	0.149	0.086	0.046	0.023	0.011	0.005	0.002	0.001	0	0	0	0	0	0
2	0.988	0.93	0.82	0.678	0.526	0.383	0.262	0.167	0.1	0.055	0.027	0.012	0.005	0.002	0	0	0	0	0
3	0.999	0.987	0.95	0.879	0.776	0.65	0.514	0.382	0.266	0.172	0.102	0.055	0.026	0.011	0.004	0.001	0	0	0
4	1	0.998	0.99	0.967	0.922	0.85	0.751	0.633	0.504	0.377	0.262	0.166	0.095	0.047	0.02	0.006	0.001	0	0
5	1	1	0.999	0.994	0.98	0.953	0.905	0.834	0.738	0.623	0.496	0.367	0.249	0.15	0.078	0.033	0.01	0.002	0
6	1	1	1	0.999	0.996	0.989	0.974	0.945	0.898	0.828	0.734	0.618	0.486	0.35	0.224	0.121	0.05	0.013	0.001
7	1	1	1	1	1	0.998	0.995	0.988	0.973	0.945	0.9	0.833	0.738	0.617	0.474	0.322	0.18	0.07	0.012
8	1	1	1	1	1	1	0.999	0.998	0.995	0.989	0.977	0.954	0.914	0.851	0.756	0.624	0.456	0.264	0.086
9	1	1	1	1	1	1	1	1	1	0.999	0.997	0.994	0.987	0.972	0.944	0.893	0.803	0.651	0.401
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Compare the value in the previous table with the **Problem 1** solved and check for the answer

N = Total number of trials **(10 in our case)**

X = Number of success we are interested **(0,1,2)**

P = Probability of success **(0.3 in our case)**

Binomial Distribution - **Ex Problem**

BINOMIAL PROBABILITY FUNCTION

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)} \tag{5.8}$$

where

$f(x)$ = the probability of x successes in n trials
n = the number of trials
$\binom{n}{x} = \dfrac{n!}{x!(n-x)!}$
p = the probability of a success on any one trial
$1 - p$ = the probability of a failure on any one trial

In the Martin Clothing Store example, let us compute the probability that no customer makes a purchase, exactly one customer makes a purchase, exactly two customers make a purchase, and all three customers make a purchase. The calculations are summarized in Table 5.7, which gives the probability distribution of the number of customers making a purchase. Figure 5.4 is a graph of this probability distribution.

The binomial probability function can be applied to *any* binomial experiment. If we are satisfied that a situation demonstrates the properties of a binomial experiment and if we know the values of n and p, we can use equation (5.8) to compute the probability of x successes in the n trials.

TABLE 5.7 PROBABILITY DISTRIBUTION FOR THE NUMBER OF CUSTOMERS MAKING A PURCHASE

x	$f(x)$
0	$\frac{3!}{0!3!}(.30)^0(.70)^3 = .343$
1	$\frac{3!}{1!2!}(.30)^1(.70)^2 = .441$
2	$\frac{3!}{2!1!}(.30)^2(.70)^1 = .189$
3	$\frac{3!}{3!0!}(.30)^3(.70)^0 = .027$
	1.000

Poisson Distribution - **Ex Problem**

($x = 0, 1, 2, \ldots$).

An Example Involving Time Intervals

Suppose that we are interested in the number of arrivals at the drive-up teller window of a bank during a 15-minute period on weekday mornings. If we can assume that the probability of a car arriving is the same for any two time periods of equal length and that the arrival or nonarrival of a car in any time period is independent of the arrival or nonarrival in any other time period, the Poisson probability function is applicable. Suppose these assumptions are satisfied and an analysis of historical data shows that the average number of cars arriving in a 15-minute period of time is 10; in this case, the following probability function applies.

$$f(x) = \frac{10^x e^{-10}}{x!}$$

The random variable here is x = number of cars arriving in any 15-minute period.

If management wanted to know the probability of exactly five arrivals in 15 minutes, we would set $x = 5$ and thus obtain

$$\begin{array}{l}\text{Probability of exactly} \\ \text{5 arrivals in 15 minutes}\end{array} = f(5) = \frac{10^5 e^{-10}}{5!} = .0378$$

Although this probability was determined by evaluating the probability function with $\mu = 10$ and $x = 5$, it is often easier to refer to a table for the Poisson distribution. The table provides probabilities for specific values of x and μ. We included such a table as Table 7 of Appendix B. For convenience, we reproduced a portion of this table as Table 5.9. Note that to use the table of Poisson probabilities, we need know only the values of x and μ. From Table 5.9

Exponential Distribution - **Ex Problem**

EXPONENTIAL DISTRIBUTION: CUMULATIVE PROBABILITIES

$$P(x \le x_0) = 1 - e^{-x_0/\mu} \tag{6.5}$$

For the Schips loading dock example, x = loading time in minutes and μ = 15 minutes. Using equation (6.5)

$$P(x \le x_0) = 1 - e^{-x_0/15}$$

Hence, the probability that loading a truck will take 6 minutes or less is

$$P(x \le 6) = 1 - e^{-6/15} = .3297$$

Relationship Between the Poisson and Exponential Distributions

In Section 5.5 we introduced the Poisson distribution as a discrete probability distribution that is often useful in examining the number of occurrences of an event over a specified interval of time or space. Recall that the Poisson probability function is

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

where

μ = expected value or mean number of occurrences over a specified interval

The continuous exponential probability distribution is related to the discrete Poisson distribution. If the Poisson distribution provides an appropriate description of the number of occurrences per interval, the exponential distribution provides a description of the length of the interval between occurrences.

To illustrate this relationship, suppose the number of cars that arrive at a car wash during one hour is described by a Poisson probability distribution with a mean of 10 cars per hour. The Poisson probability function that gives the probability of x arrivals per hour is

$$f(x) = \frac{10^x e^{-10}}{x!}$$

Because the average number of arrivals is 10 cars per hour, the average time between cars arriving is

$$\frac{1 \text{ hour}}{10 \text{ cars}} = .1 \text{ hour/car}$$

Thus, the corresponding exponential distribution that describes the time between the arrivals has a mean of $\mu = .1$ hour per car; as a result, the appropriate exponential probability density function is

$$f(x) = \frac{1}{.1} e^{-x/.1} = 10e^{-10x}$$