UNIT 2

Bayes Theorem, Central Limit Theorem

Presented By :

Aditya Joshi

Assistant Professor

Graphic Era Deemed to be University, Dehradun

Conditional Probability

- In conditional probability we have to find out the probability of something given, that we have some other piece of information.

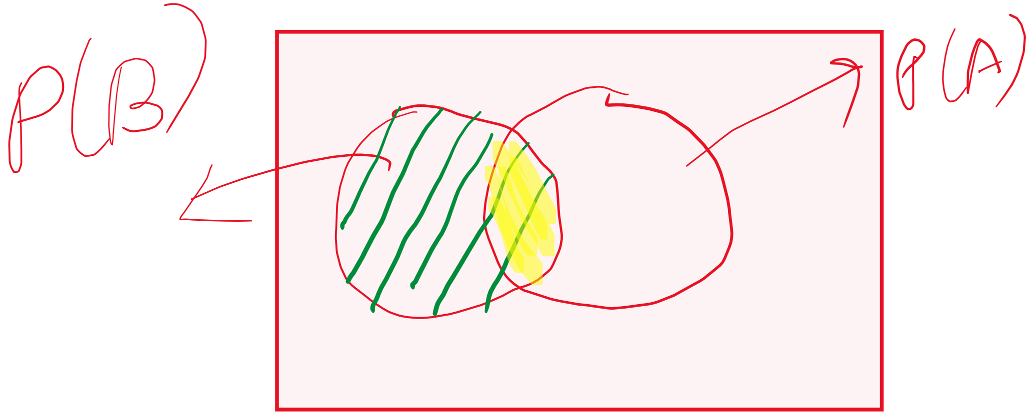
$P(A|B)$ = The Probability of A, the thing I want the thing I am investigating, GIVEN that I have this other piece of information given that I Know that event B has occurred

given

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

Intersection

Probability of Both the event A & the event B



- The % of adults who are men and alcoholics is 2.25%. And prob of being a man is 50% What is the probability of being an alcoholic , given being a man ?

$$P(\text{Alcholalic} \mid \text{man}) = \frac{P(\text{Alcholic} \cap \text{man})}{P(\text{man})}$$

$$= \frac{0.0225}{.5}$$

$$= \underline{0.045}$$

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

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

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

There are two children of a parent.What is the probability of two girls given at least one girl?

$$P(\text{2 Girls}|\text{at least 1 girl}) = \frac{P(\text{1 Girl}|\text{2 Girls}) * P(\text{2 Girl})}{P(1\, Girl)}$$

GG BG GB BB

$$= \frac{1 \times \frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

Bayesian Modeling

- Bayesian Model is a probabilistic model(a system of making inference) that is based on bayes theorem. The Bayesian Model attempts to obtain a posterior distribution. Given by Thomas Bayes.

For example, if we have the density function for some observations X_i

For i=1 to n to be $f(X_i | \theta)$ for unknown parameter θ

Then the Prior distribution is given by p(θ), Bayesian Model would try to find out the parameter using the posterior $p(\theta | X)$

A function that relates one variable to another

Bayes Theorem is

Posterior Probability

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Prior Probability

$$P(Y/X) = [K]\, P(Y)$$

$$K = \frac{P(X|Y)}{P(X)}$$

x	y
2	4
3	6
4	8
5	10
x_i	?

Bayesian modeling

$$f(x) = y$$

Regression

$$f(x) = \beta_0 + \beta_1 x$$

$$y = mx + c$$

$$y = a + bx$$

Population → Samples n $x_1, x_2, x_3 \ldots x_n$

θ parameter

likelyhood

Prior Prob.

Posterior Probability ← $P(\theta \mid x_1, x_2, x_3 \ldots x_n) = \dfrac{P(x_1, x_2, x_3 \ldots x_n \mid \theta)\, P(\theta)}{P(x_1, x_2, x_3 \ldots x_n)}$

Marginal Probability

OR

Marginal Prior

Central Limit Theorem

The Central Limit Theorem states that the sampling distribution of the sample means approaches a normal distribution **as the sample size gets larger** — no matter what the shape of the population distribution. This fact holds especially true for sample sizes over 30.

All this is saying is that as you take more samples, especially large ones, your graph of the sample means will look more like a **normal distribution**.

Here’s what the Central Limit Theorem is saying, graphically. The picture below shows one of the simplest types of test: rolling a fair die. The more times you roll the die, the more likely the shape of the distribution of the means tends to look like a normal distribution graph.

A Central Limit Theorem word problem will most likely contain the phrase "assume the variable is normally distributed", or one like it. With these central limit theorem examples, you will be given:

A population (i.e. 29-year-old males, seniors between 72 and 76, all registered vehicles, all cat owners)

An average (i.e. 125 pounds, 24 hours, 15 years, $15.74)

A standard deviation (i.e. 14.4lbs, 3 hours, 120 months, $196.42)

A sample size (i.e. 15 males, 10 seniors, 79 cars, 100 households)

You will get one of three central limit theorem examples:

I want to find the probability that the mean is **greater** than a certain number

I want to find the probability that the mean is **less** than a certain number

I want to find the probability that the mean is **between** a certain set of numbers either side of the mean