Conditional Probability and Bayes's Theorem

Content

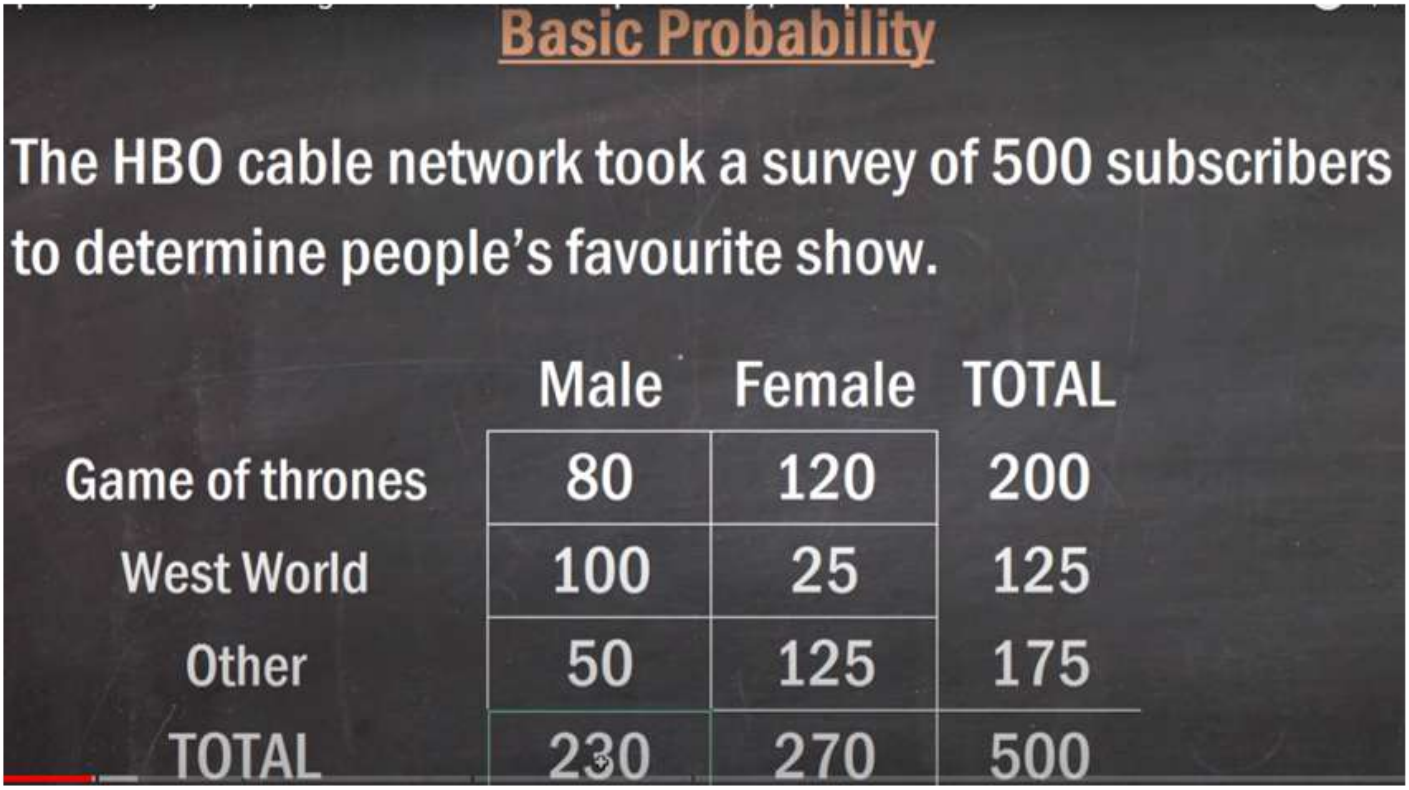
- Sample Spaces and Events
- Joint, Conditional and Marginal Probability
- Bayes's Theorem

Probability Theory – Terminology..

- **Event-**a subset of a sample space and probability is usually calculated with respect to an event. **For Example:**

    1. Number of cancellation of orders placed at an E-commerce portal site exceeding 10%
    2. The number of fraudulent credit card transactions exceeding 1%
    3. The life of a capital equipment being less than one year

Probabilities: Joint, Conditional and marginal

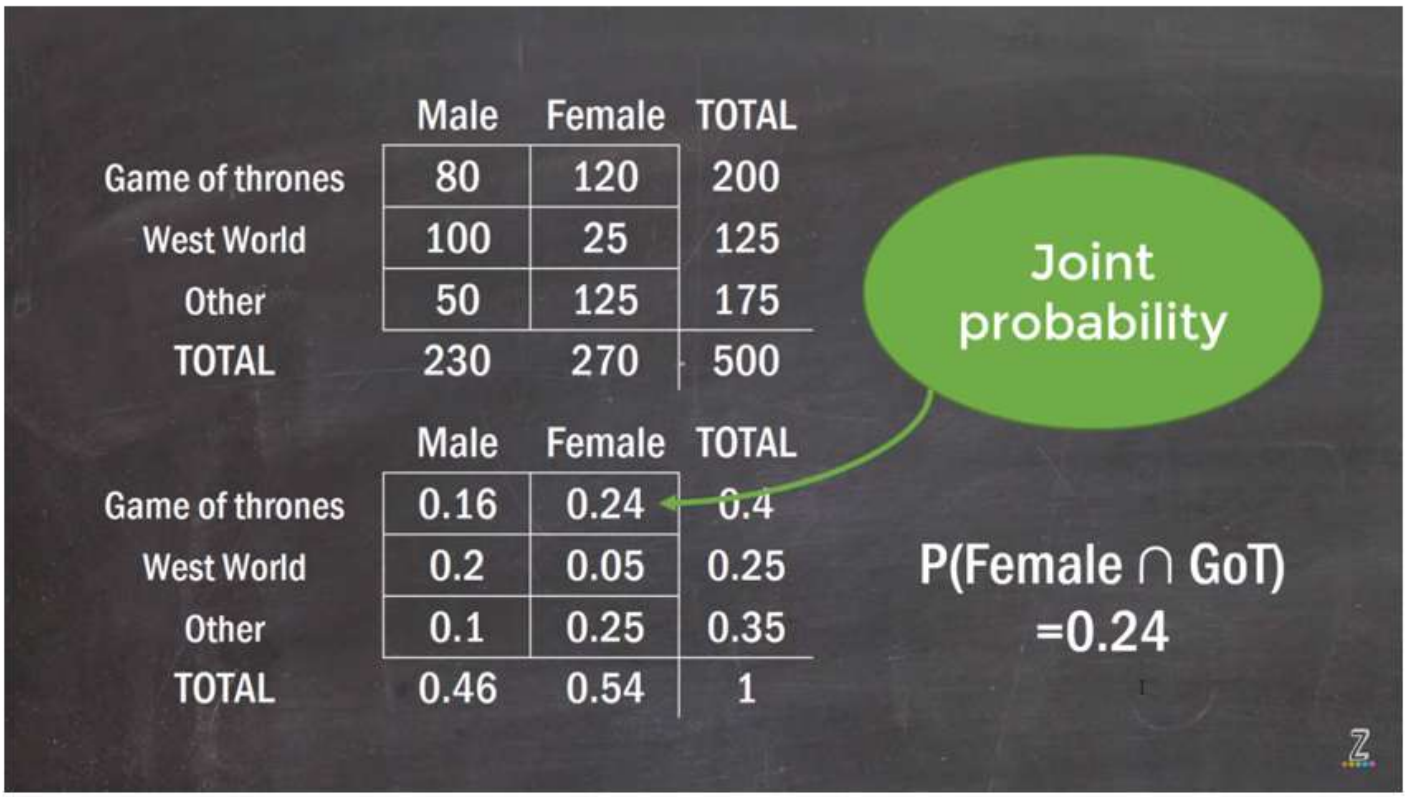


Basic Probability

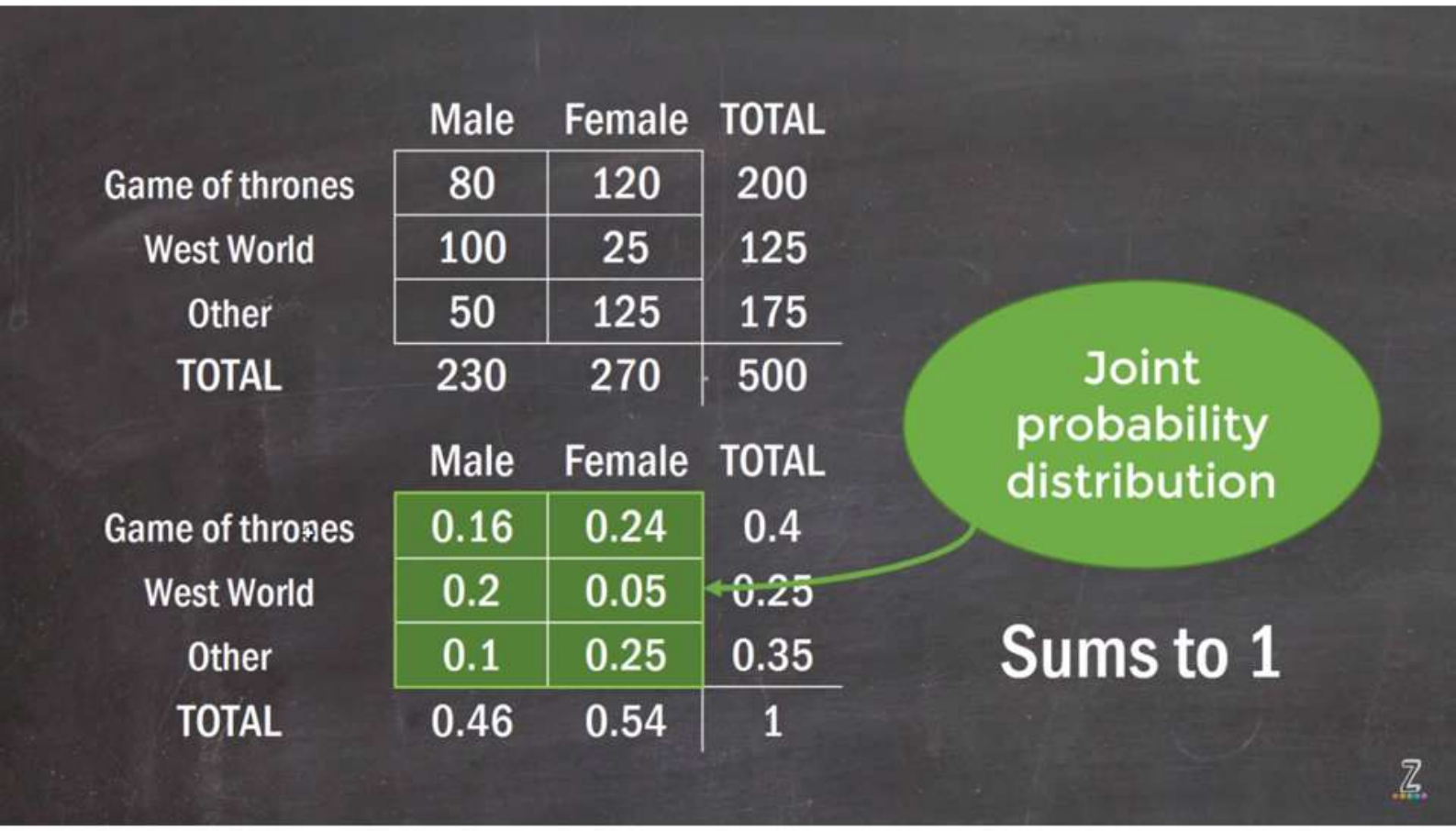
The HBO cable network took a survey of 500 subscribers to determine people's favourite show.

	Male	Female	TOTAL
Game of thrones	80	120	200
West World	100	25	125
Other	50	125	175
TOTAL	230	270	500

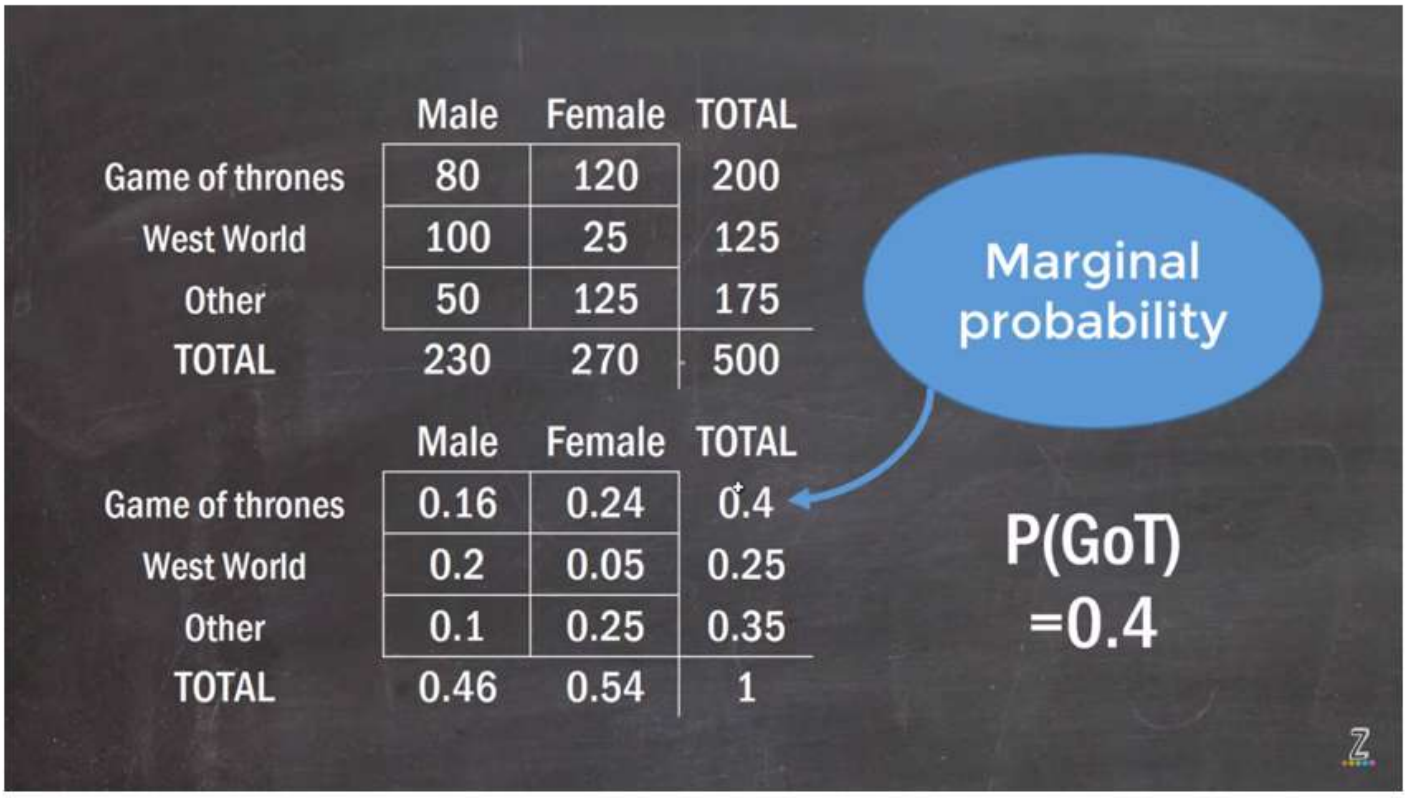
Joint Probability: To get divide everything by 500



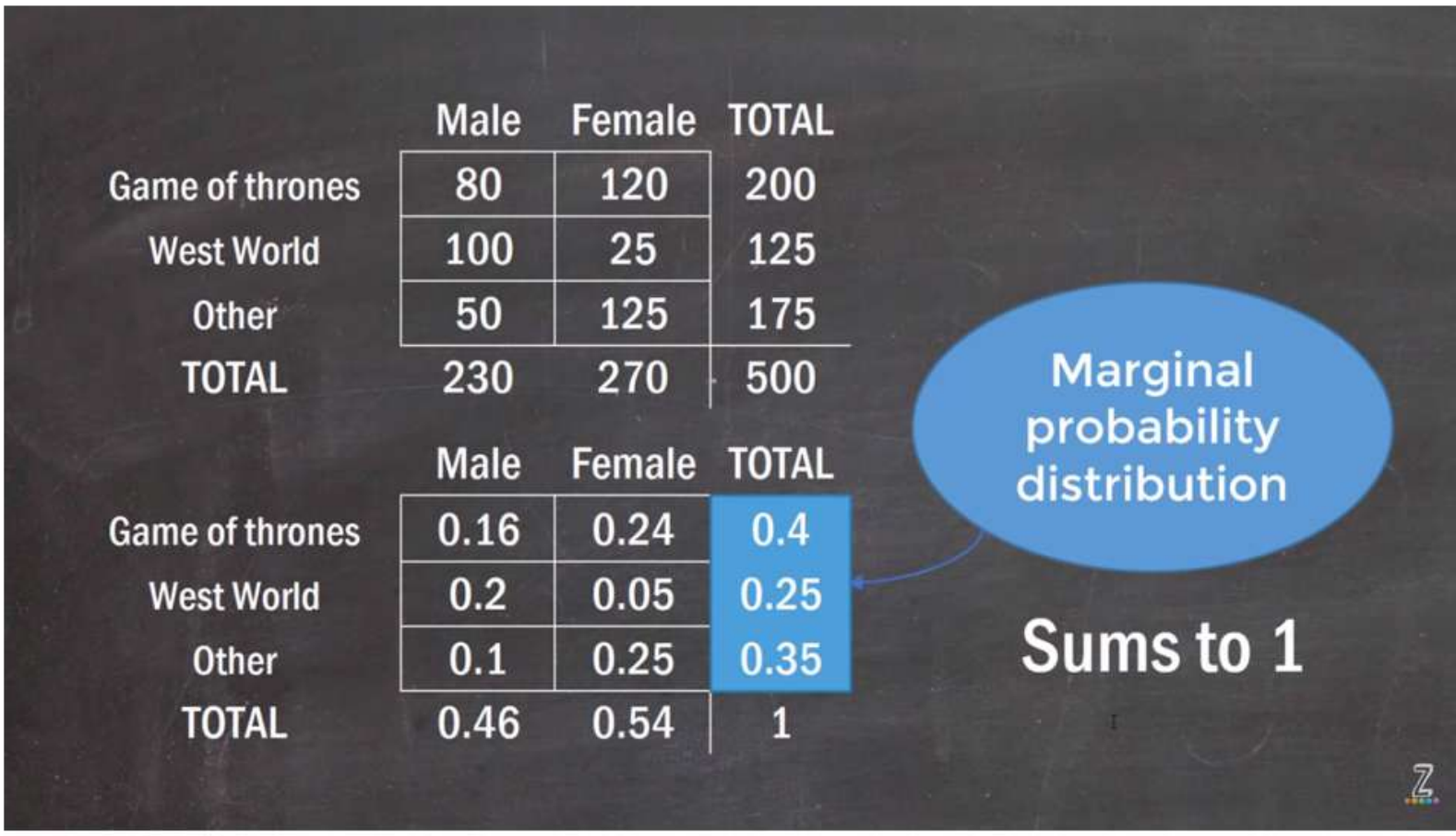
Joint Probability...



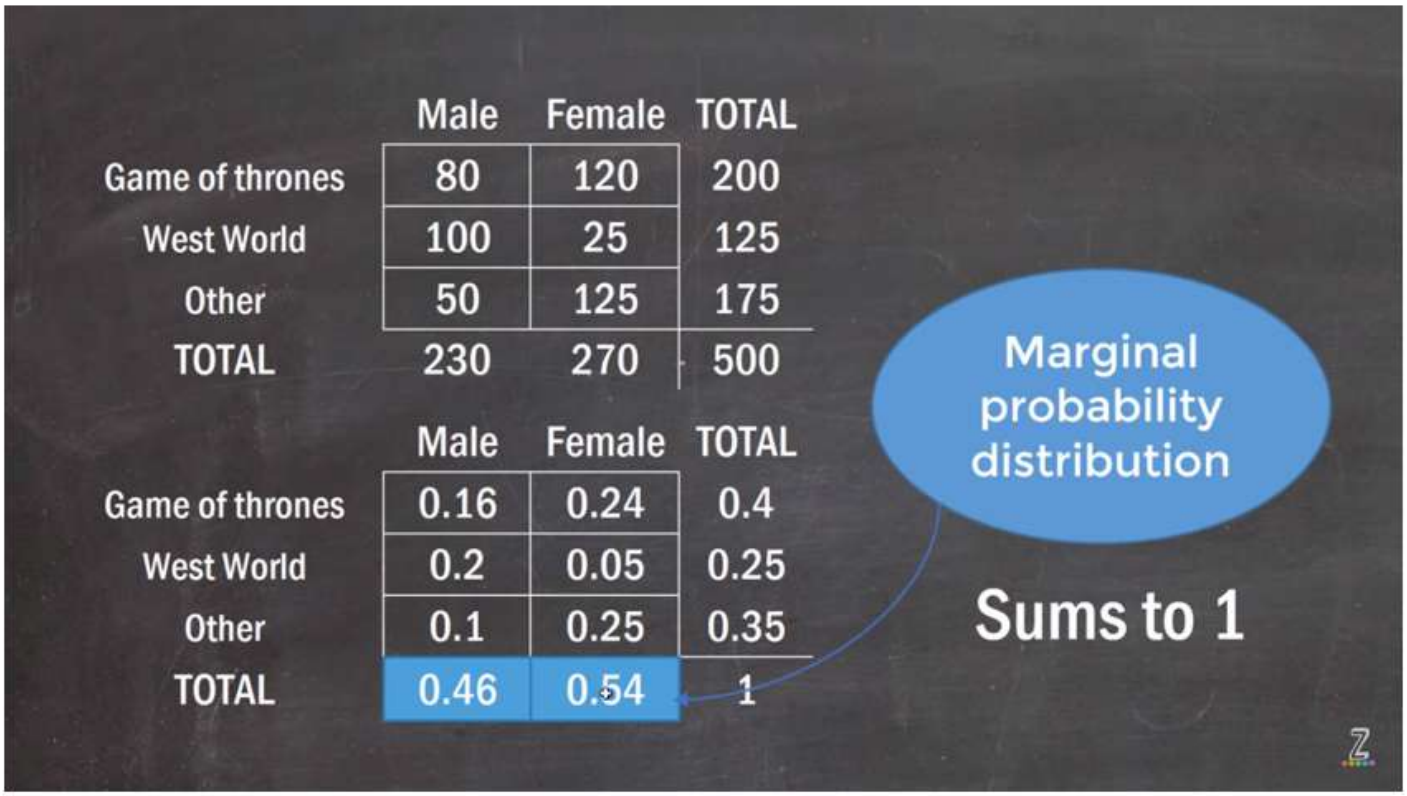
Marginal Probability: Row Probability



Marginal Probability...



Marginal Probability...



Questions on Probability...Q1

	Male	Female	TOTAL
Game of thrones	0.16	0.24	0.4
West World	0.2	0.05	0.25
Other	0.1	0.25	0.35
TOTAL	0.46	0.54	1

Q: What is the probability of an HBO subscriber being male?

P(Male) = 0.46

Questions on Probability...Q2

	Male	Female	TOTAL
Game of thrones	0.16	0.24	0.4
West World	0.2	0.05	0.25
Other	0.1	0.25	0.35
TOTAL	0.46	0.54	1

Q: What is the probability of an HBO subscriber preferring West World?

P(West World) = 0.25

Questions on Probability...Q3

	Male	Female	TOTAL
Game of thrones	0.16	0.24	0.4
West World	0.2	0.05	0.25
Other	0.1	0.25	0.35
TOTAL	0.46	0.54	1

Q: What is the probability of an HBO subscriber being male **AND** preferring West World?

$$P(\text{Male} \cap \text{West World}) = 0.2$$

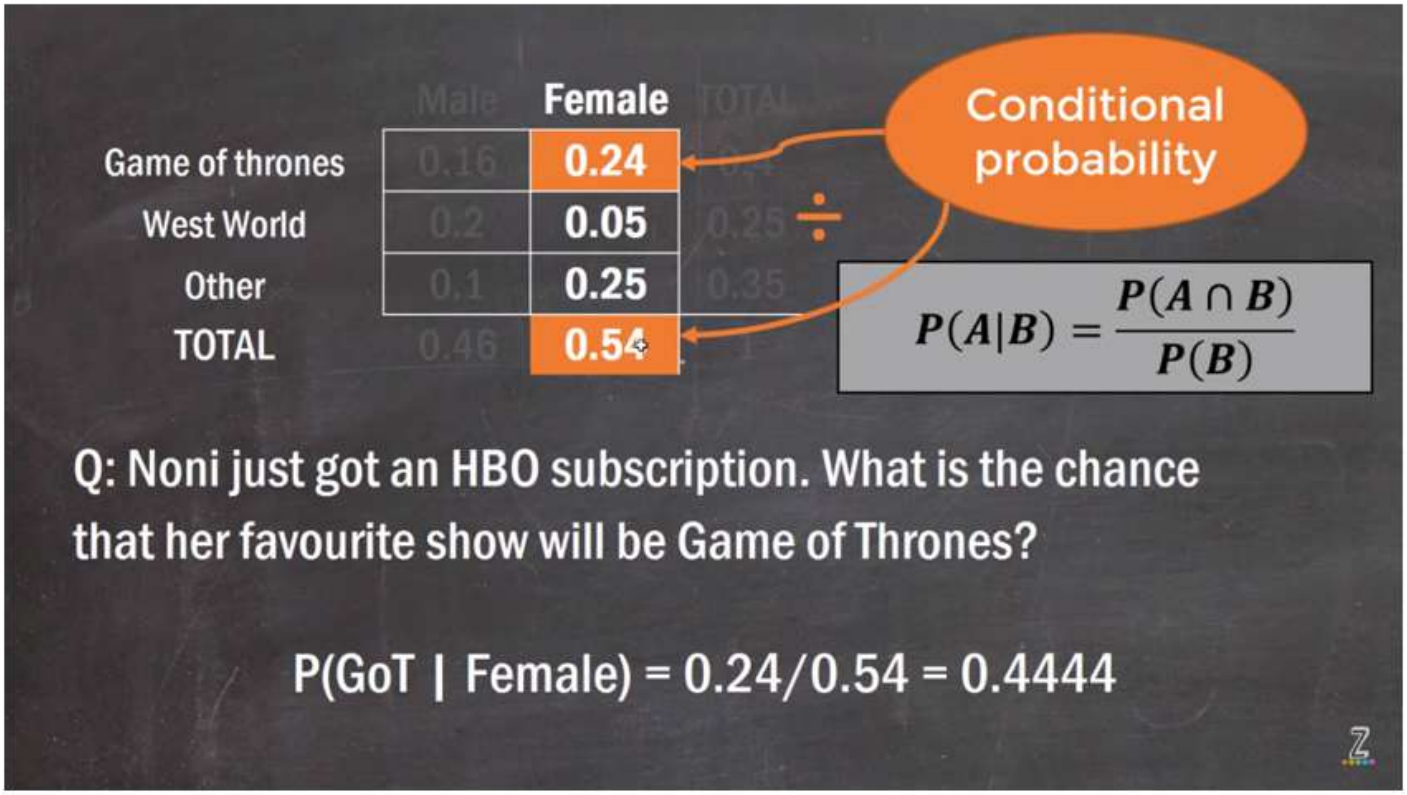
Questions on Probability...Q4

	Male	Female	TOTAL
Game of thrones	0.16	0.24	0.4
West World	0.2	0.05	0.25
Other	0.1	0.25	0.35
TOTAL	0.46	0.54	1

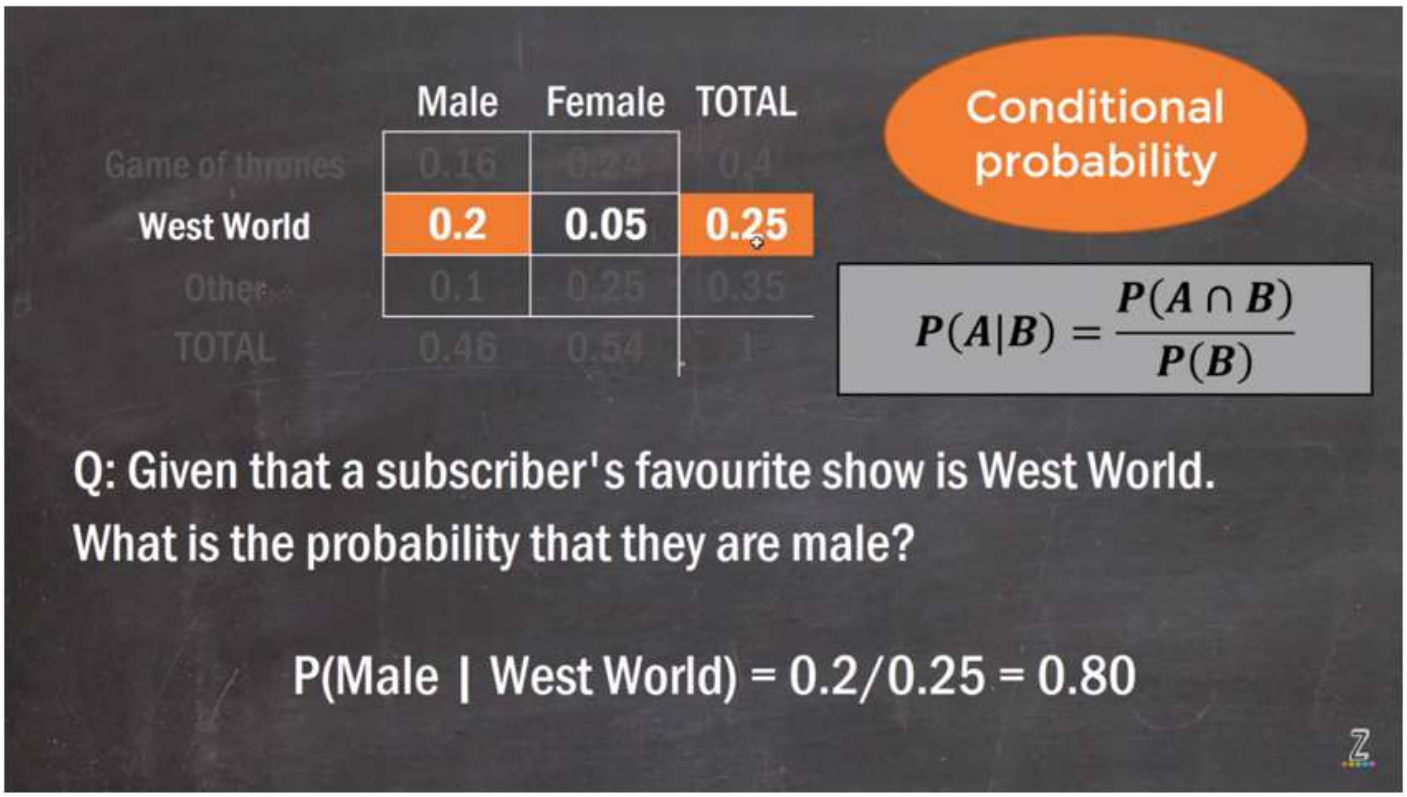
Q: What is the probability of an HBO subscriber being male OR preferring West World?

$$P(\text{Male} \cup \text{West World}) = 0.16+0.2+0.1+0.05$$
$$= 0.51$$

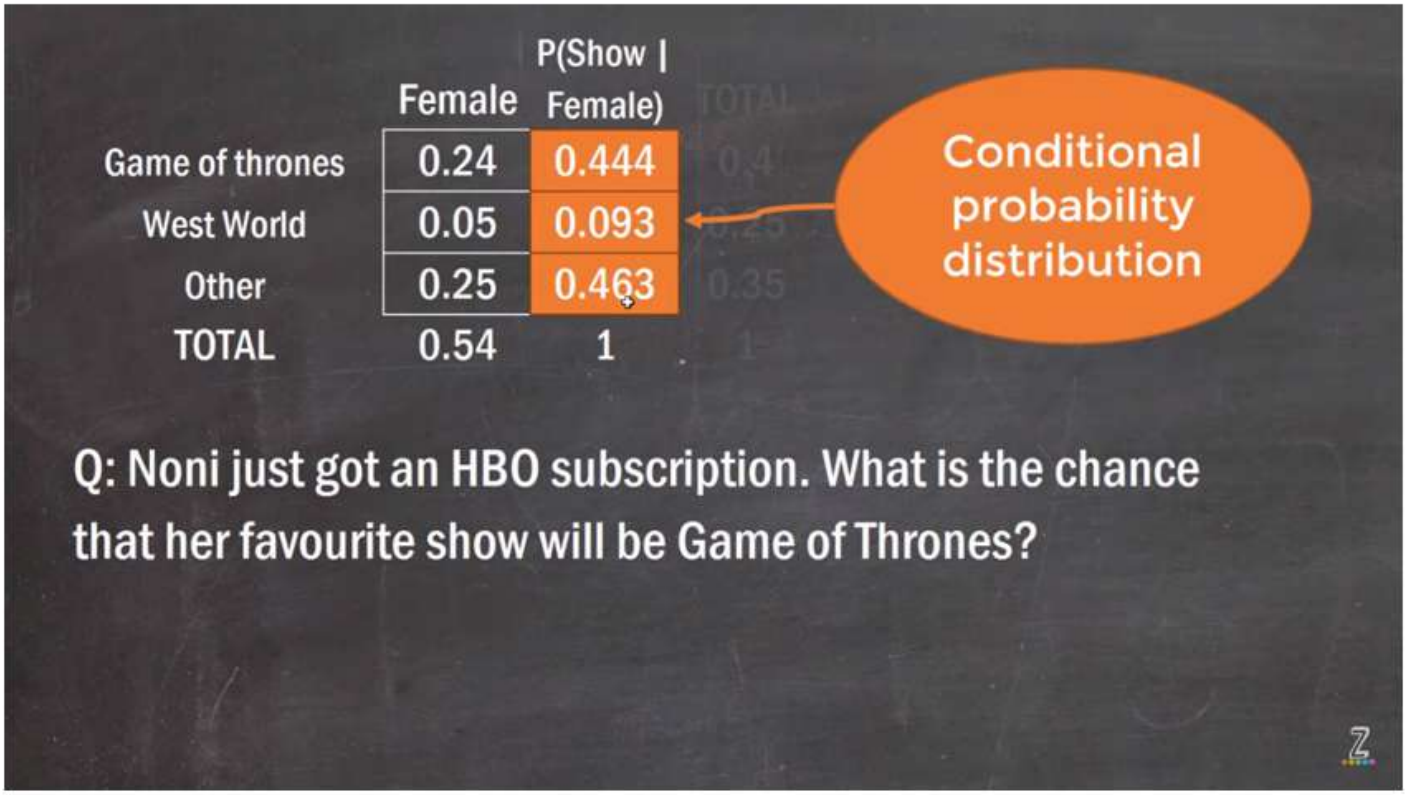
Conditional Probability



Conditional Probability..



Conditional Probability distribution



Naïve Bayes. Theorem

Naïve Bayes Classifier

- Probabilistic classifier
- Widely used in Text categorization

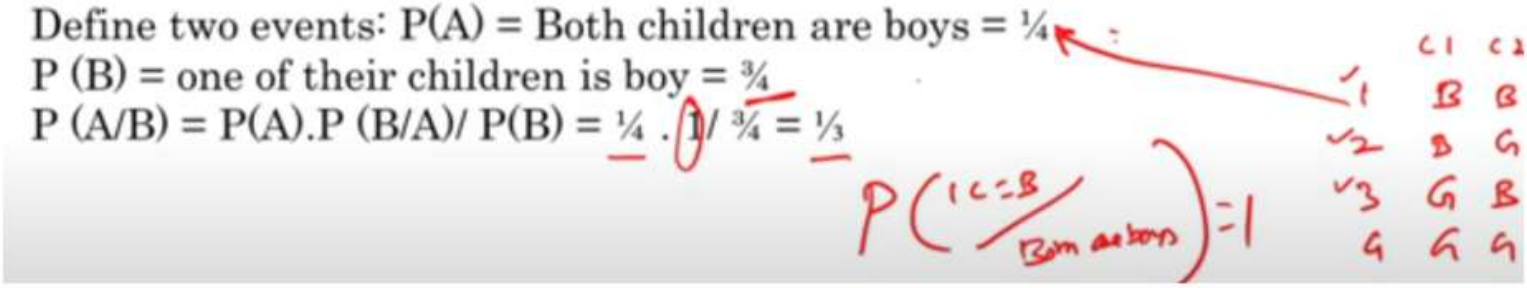
- Supervised algorithm
- Bayes Theorem
- Naïve assumption – occurrence of features are independent
  - If we model students performance based on attendance, assignment submission. Assumption is occurrence of assignment submission and attendance are independent

Note: That is P(A and B) is not accepted. It is his assumption

Question on Naïve Bayes

A couple has two children, one of which is a boy. What is the probability that they have two boys?

Define two events: P(A) = Both children are boys = ¼
P (B) = one of their children is boy = ¾
P (A/B) = P(A).P (B/A)/ P(B) = ¼ . 1/ ¾ = ⅓



Lab activity

- We load the Iris dataset from scikit-learn.
- Split the dataset into training and testing sets.
- Initialize a Gaussian Naive Bayes classifier (GaussianNB).
- Train the classifier using the training data.
- Make predictions on the testing data using the trained classifier.
- Evaluate the accuracy of the classifier by comparing the predicted labels with the actual labels from the testing set.