Day 2 Notes - Probability

# Conditional Probability

Walk through these examples:

<http://www.stat.yale.edu/Courses/1997-98/101/condprob.htm>

<https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/conditional-probability-definition-examples/>

### **Examples**

In a card game, suppose a player needs to draw two cards of the same suit in order to win. Of the 52 cards, there are 13 cards in each suit. Suppose first the player draws a heart. Now the player wishes to draw a second heart. Since one heart has already been chosen, there are now 12 hearts remaining in a deck of 51 cards. So the conditional probability *P(Draw second heart|First card a heart)* = 12/51.

Suppose an individual applying to a college determines that he has an 80% chance of being accepted, and he knows that dormitory housing will only be provided for 60% of all of the accepted students. The chance of the student being accepted *and* receiving dormitory housing is defined by

*P(Accepted and Dormitory Housing) = P(Dormitory Housing|Accepted) \* P(Accepted)* = (0.60)\*(0.80) = 0.48.

**Combined probability exercise**

* In a 52 deck card, what is the chance of getting a 'picture card' (king, queen, jack)
* What is the probability of getting a RED card
* What is the combined probability of getting a RED and picture card?
* What is the probability of getting a diamond card
* What is the probability of getting PICTURE card given DIAMOND card

**Answers**

Probability (picture card) = 4 kings + 4 queens + 4 jacks = 12 / 52

P (red) = 26 / 52 = 50%

P (red and picture) = P (picture) \* p (red) = (12/52) \* (26/52) = 0.12

P (diamond) = 13 / 52

P (picture | diamond) = P (picture AND diamond) / P (diamond) = (3/52) / (13/52) = 0.23



# Histogram

<https://en.wikipedia.org/wiki/Freedman%E2%80%93Diaconis_rule>

# Distributions

## Bernoulli Distribution

<https://www.statisticshowto.com/bernoulli-distribution/>

<https://www.youtube.com/watch?v=bT1p5tJwn_0>

## Binomial distribution

<https://www.statisticshowto.com/probability-and-statistics/binomial-theorem/binomial-distribution-formula/>

<https://www.youtube.com/watch?v=qIzC1-9PwQo>

Exercises

[Nice worksheet with python examples](https://projector-video-pdf-converter.datacamp.com/14568/chapter1.pdf) - work through this one!

## Poisson Distribution

<https://www.statisticshowto.com/poisson-distribution/>

<https://www.youtube.com/watch?v=jmqZG6roVqU>

<https://towardsdatascience.com/the-poisson-distribution-and-poisson-process-explained-4e2cb17d459>

## Confidence intervals

<https://www.simplypsychology.org/confidence-interval.html>

## Z Score calculations

<https://www.youtube.com/watch?v=sJyZ9vRhP7o>

<http://www.z-table.com/>

# Large Scale ML

<https://spark.apache.org/docs/latest/ml-guide.html>

<https://www.h2o.ai/>

<https://aws.amazon.com/sagemaker/>

## More Exercises

<https://medium.com/analytics-vidhya/probability-statistics-for-beginners-in-machine-learning-part-3-probability-distribution-9bce0567fdcf>

<https://elephantscale.github.io/guided-machine-learning/python-data-analysis/stats-1__basic-stats.html>