**Addition Rule**

Union of any event P(A or B) = P(A) + P(B) – P(A and B), where P(A andB) = 0 if they are mutually exclusive

**Product Rule**

P(A and B) = P(A|B) \* P(B)

if P(A|B) = P(A), then A and B are independent

**Bayes Theorem**

, if P(A|B) = P(A), then A and B are independent

**Application of Bayes Thorem**

**Binomial Dist**

**in R =** **dist <- dbinom(k, n, p)**

P(binom) = # of scenarios\*P(single scenario)= \*

mean of binomial dist µ=np

st dev for binom dist σ=√(np(1-p))**Normal Dist**

68-95-99.7% rule for SD

**Z-score**

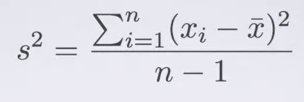
**aka in R z <- pnorm(x, mean, st dev)**

zscore of mean = 0, unusual observations are given as |Z| >2

**Find Cutoffs using Z**

**R score\_cutoff <- qnorm(percentile, mean, st dv)**

**Variance/ SD**

, where the ST Dev is the square root of that

**Posterior Probability**

Defined as P(Hypothesis|data observed), Different than the **pValue which is P(data|hypothesis)**

Generally, the probability of at least one is the same as 1 –P(none)

Quantile Plot in R (if it follows y=x, then it’s nearly normal)

1. x = vector
2. >
3. n = length(x)
4. >
5. plot((1:n - 1)/(n - 1), sort(x), type="l",
6. main = "Quantiles for the NYC Rain Data",
7. xlab = "Sample Fraction",
8. ylab = "Sample Quantile")

**Central Limit Theorem**

The Central Limit Theorem says that if you

• take many random samples of size n

• from a population with true proportion = p

then the many sample proportions will be

* approximately normally distributed, with

• mean equal to p, and

* + standard deviation equal to the square root of

p \* (1 – p) / n

If a sample consists of at least 30 independent observations and the data are not

strongly skewed, then the distribution of the sample mean is well approximated by

a normal model.

The distribution of \_x is approximately normal. The approximation can be poor if

the sample size is small, but it improves with larger sample sizes.