Probability

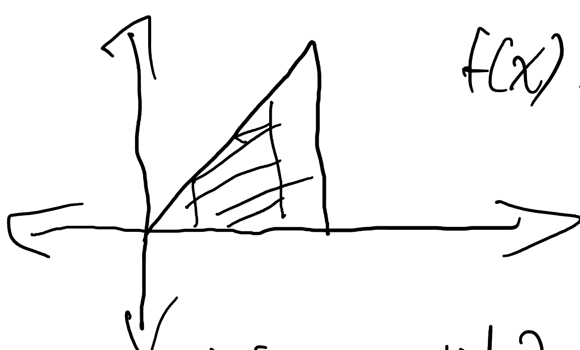
Saturday, August 13, 2016 2:32 PM

- Probability calculus helps understand rules
- Very complicated though, simplify it via **density** and **mass** functions
- Random Variables- numerical outcome of experiment
    - Discrete
    - Continuous
- Probability Mass Function:
    - Discrete
    - Function that associates every outcome with a probability
    - Must always be >= 0
    - Sum of possible vars must always = 1
- Prevalent example:
    - $(\theta)^x (1-\theta)^{1-x} = p(\theta, x)$

      θ = prob. of event

      x = ~~#~~ prob. of trials

- Probability Density Function:
    - Continuous
    - Must be larger than 0 everywhere
    - Total area under it must be 1
    - Areas under PDFs correspond to prob for that random var



$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \end{cases}$$

Is this a valid?

Area = 1

Always ≥ 0

Cumulative Distribution Function

Probability var $\leq x$

$$F(x) = P(X \leq x)$$

Survival Function:

Opposite of CDF

$$S(x) = P(X > x)$$

Quantile:

The α^{th} quantile $\leftrightarrow$ percentile

CDF $\rightarrow$ $F(x) = \alpha$

Conditional Probability

Saturday, August 13, 2016 3:15 PM

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

Baye's Rule:

$P(B|A)$ if you know $P(A|B)$

Diagnostic Tests:

$+$ and $-$ events result of diagnostic test

D and D^c event that subject has or does not have disease

retrievable

$P(+|D)$ = sensitivity

$P(-|D^c)$ = specificity

What people care about:

$P(D|+)$

convert

$P(D^c|-)$

Baye's Formula:

$P(A|B)P(B)$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

Likelihood Ratios:

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

Independence

Saturday, August 13, 2016 3:41 PM

- A is independent of B if:
    - P(A|B) = P(A) where P(B) > 0
    - P(A^B) = P(A)P(B)
- Independent
    - Statistically independent of another event
- Identical Distribution
    - Drawn from the same population distribution

Expected Values

Saturday, August 13, 2016 3:46 PM

- Making conclusions based on noisy data
- Sample expected values will predict population expected values
  - Mean
    - $E[X] = \sum_x xp(x)$
  - Variance
- Population mean is center of mass of population
- Sample mean is center of mass of sample population
- Sample mean is estimate of population mean
- More data in sample mean, closer to actual mean it is

Variability

Saturday, August 13, 2016 4:04 PM

- Variance
  - Variance of random var is measure of spread
  - Var(X) = E[(X-u)]^2
  - Square root of variance is std deviation
  - Sample Variance:

$$S^2 = \frac{\sum_{i-1}(X_i - \bar{X})^2}{n-1}$$

Screen clipping taken: 8/13/2016 4:07 PM

- Sample Variance
  - Distribution of sample variance centered at what's being estimated
  - Gets more concentrated and thus closer to population variance with more samples
  - Variance of sample mean is population variance divided by n

Binomial Distribution

Saturday, August 13, 2016 4:31 PM

- Bernoulli Distribution
    - Binomial is a series of Bernoullis
- $(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k}$

Normal Distribution

Saturday, August 13, 2016 4:41 PM

- Centered on mean
- Standard deviations have equal number of values
- 1 std dev away- 68% of data
- 2 std devs- 95%
- 3 std devs- 99%

Poisson Distribution

Saturday, August 13, 2016 4:44 PM

- Used to model counts

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

Screen clipping taken: 8/13/2016 4:45 PM

- Mean and variance are lambda
- Uses:
  - Count data (especially unbounded)
  - Unknown bounds
  - Contingency tables (qualitative values basically quantified)
  - Approximating binomials if n is too large and p is too small
- Poisson random vars represent rates
  - $X \sim Poisson(\lambda t)$
    - $\lambda = E\left[\frac{X}{t}\right]$ is the expected count per unit time
    - t is total monitoring time

Asymptotics and LLN

Sunday, August 14, 2016 12:11 PM

- Asymptotics
    - Behavior of estimators as sample size approaches infinity
    - Frequency interpretation of probabilities
- Law of Large Numbers
    - Average limits what its estimating, the population mean
    - As number of samples goes up, converge to true value
- Estimator is **consistent** if converges to what you want to estimate
- Good estimators must be consitent, with consistent variance and mean

Central Limit Theorem

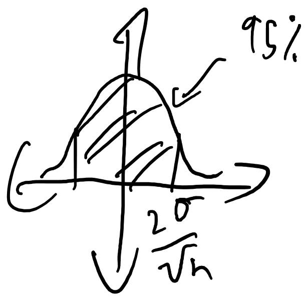
Sunday, August 14, 2016 12:19 PM

- Central Limit Theorem
  - Distribution of averages of iid (Independent Identically Distributed) (properly normalized) approaches standard normal as sample size increases

Confidence Intervals

Sunday, August 14, 2016 12:28 PM

$\bar{X}$ is $\simeq$ normal w/ mean μ and sd $\frac{\sigma}{\sqrt{n}}$



$\bar{X} \pm \frac{2\sigma}{\sqrt{n}}$ is 95% confidence interval

For examples, look at Mike's notes