Basics **Mean** $\bar{x} = \frac{\sum x_i}{n}$

Median - n **odd** $\tilde{x} = \left(\frac{n+1}{2}\right)^{th}$ value

Median - n **even** Average of $\left(\frac{n}{2}\right)^{th}$ and $\left(\frac{n}{2}+1\right)^{th}$ values

Variance $s^2 = \sigma^2 = \frac{\sum(x_i - \bar{x})}{n-1}$

Quartiles **Fourths** Upper fourth = median of upper half, etc. $\tilde{x}$ in both halfves if n odd

Fourth Spread $f_s = f_{upper} - f_{lower}$

Outlier if $x_i > 1.5 f_s$ away from closest fourth, extreme if $> 3 f_s$

Probability **Events:** An event is simple if it consists of one outcome, complex if more than one

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

independence iff $P(A|B) = P(A)$ and iff $P(A \cap B) = P(A) \cdot P(B)$

Props: For any two events $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Bayes Thrm: Let A_1,...,A_k be mutually exclusive and exhaustive events. Then for any other event B, $\sum P(B|A_i)P(A_i)$

Combinatorics **Choose:** $\binom{n}{r} = \frac{n!}{k!(n-k)!}$

Permute: $P_{k,n} = \frac{n!}{(n-k)!}$

DRV basics **Bernoulli rv:** Any rv whose only possible values are 0 and 1

pmf $p(x) = P(X = x) = P(s \in S : X(s) = x)$

cdf $F(x) = P(X \leq x) = \sum p(y)$ for $y \leq x$

Expected Value $E(x) = \mu_x = \sum x \cdot p(x)$

For a function: $E[h(x)] = \sum h(x) \cdot p(x)$

Variance: $V(X) = \mu_x^2 = \sum (x - \mu)^2 \cdot p(x) = E[(X - \mu)^2]$

Binomial Probability Distribution Sampling w/o replacement from a finite dichotomous pop. Assumes $n << N$

Fits if:

1. Some number of trials, n, where n is fixed
2. Each trial can result in either Success(S) or Failure(F)
3. Each trial is independent
4. $P(S)$ is constant

binpmf $b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$ if $x \in 0, 1, 2, \ldots, n$

bincdf $B(x; n, p) = P(X \leq x) = \sum b(y; n, p) for y \in [0, n]$

mean $E(X) = np$

variance $V(X) = np(1-p)s$

Hypergeometric Bin. Distributions **Assumptions:**

1. Finite population - N
2. Each individual either a success or failure, M successes
3. Sampled without replacement such that each subset of n is equally likely to be chosen

pmf: $P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$

where $\max(0, n - N + M) \leq x \leq \max(n, M)$

mean: $E(X) = n \cdot \frac{M}{N}$

variance: $V(X) = \frac{N-n}{N-1} \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$

Negative Binomial Distribution **Assumptions:**

1. Sequence of independent trials
2. Each trial can result in success or failure
3. The probability of success is constant
4. Trials continue until r successes

pmf: $nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x$
where p is P(success), r is # of successes, and x is # of failures preceeding the r^{th} success.

cdf:

mean: $E(X) = \frac{r(1-p)}{p}$

variance: $V(X) = \frac{r(1-p)}{p^2}$

Poisson Distribution a drv has Poisson distribution with parameter $\mu (\mu > 0)$ if the pmf is
$p(x;\mu) = \frac{e^{-\mu} \cdot \mu^x}{x!}$ where $x = 0, 1, 2, 3, \ldots$
$E(X) = V(X) = \mu$
Application: events occuring over time. Properties:

1. $\exists\ \alpha > 0,\ \Delta t > 0,\ P(\text{oneeventoccursin } \Delta t) = \alpha \cdot \Delta t + o(\Delta t)$ where $o(\Delta t) << \Delta t$
2. The probability of more than one event occurring during Δt is $o(\Delta t)$, so $P(\text{noevents}) = 1 - \alpha \cdot \Delta t - o(\Delta t)$
3. The number of events occurring in Δt is independent of the number occuring previously

So, $P_k(t) = \frac{e^{-\alpha t} \cdot (\alpha t)^k}{k!}$

Continuous Distributions Uniform on $[A, B]$ if pdf: $f(x; A, B) = \begin{cases} \frac{1}{B-A} & A \le x \le B \\ 0 & \text{otherwise} \end{cases}$

Variance $\sigma_X^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \dot{f}(x) dx = E[X - \mu^2] = E(X^2) - E(X)^2$

Normal if pdf is $f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$
Standard if $\mu = 0$, $\sigma = 1$, denoted Z
cdf denoted $\Phi(z)$

Nonstandard: Let $Z = \frac{X-\mu}{\sigma}$
Then $P(a \le x \le b) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$

Exponential $f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$
$\mu = \frac{1}{\lambda}$, $\sigma^2 = \mu^2$
cdf: $F(x; \lambda) = 1 - e^{-\lambda x}$
If # of events in a time t has Poisson distributiuon with param αt then time between events has exp. dist. with $\lambda = \alpha$

Γ function
- $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$
- $\alpha > 1 \to \Gamma(\alpha) = (\alpha - 1)\dot{\Gamma}(\alpha - 1)$
- $n \in \mathbb{N} \to \Gamma(n) = (n-1)!$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Gamma distribution pdf: $f(x; \alpha, \beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}$ where α, $\beta > 0$.

Standard pdf: if $\beta = 1$: $f(x, \alpha) = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)}$

cdf: $F(x; \alpha, \beta) = F(\frac{x}{\beta}; \alpha)$ where F is the incomplete Γ function.
$F(y; \alpha) = \int_0^y \frac{y^{a-1} e^{-y}}{\Gamma(\alpha)} dy$, assuming Y has a std. gamma dist

measures: $E(X) = \alpha\beta$, $V(X) = \alpha\beta^2$

χ^2 distribution $f(x; v)$, Γ dist. with $\alpha = v/2$, $\beta = 2$

Joint Distributions Independent iff $f(x, y) = f_X(x)\dot{f}_Y(y)$
Conditional: $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$
Expected Value: $E[h(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)\dot{f}(x, y) dx dy$
Covariance: $Cov(x, y) = E[(X - \mu_x)(Y - \mu_y)] = E(XY) - \mu_X \mu_Y$
Correlation coefficiant: $\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$

Random Samples Let $X_1, X_2 \ldots X_n$ be a random sample from a dist. with mean μ and s.d. σ
$E(\bar{X}) = \mu$, $V(\bar{X}) = \sigma_{\bar{X}}^2 = \sigma^2/n$
If dist is normal then for any n, $\bar{X}$ is norm. distributed with mean μ and s.d. $\sigma/\sqrt{n}$