Results from Probability and Statistics

- PMF (probability pass function): $p_X(x) = \Pr(X = x)$
  - For discrete RVs, satisfies $p_X(x) \geq 0$ and $\sum_x p_X(x) = 1$
- PDF (probability density function): $f_X(x)$
  - For continuous RVs, satisfies $f_X(x) \geq 0$ and $\int_x f_X(x)dx = 1$
- CDF (cumulative density function) $F_X(x) = \Pr(X \leq x)$
  - For continuous RVs, $f_X(x) = \frac{\partial}{\partial x} F_X(x)$
- Expectations:
  - $E[X] = \sum_x x p_X(x)$if X is discrete
  - $E[X] = \int_x x f_X(x)dx$ if X is continuous
  - $E[g(X)] = \sum_x g(x)p_X(x)$ if X is discrete
  - $E[g(X)] = \int_x g(x)f_X(x)dx$ if X is continuous
- Population Mean and Variance
  - $\mu_X = E[X]$
  - $\sigma_X^2 = Var(X) = E[(X - \mu_X)^2]$
  - $Var(X) = E[(X - \mu_X)^2] = E[X^2] - E[X]^2$
- Linear Transformations
  - $E[aX + b] = aE[X] + b$
  - $Var(aX + b) = a^2 Var(X)$
- Normal Distribution: $f_X(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2} / \sigma^2$
- If $X_j \sim N(\mu_j, \sigma_j^2)$ and X_j are independent, $w_1 X_1 + ... + w_J X_J \sim N(\mu, \sigma^2)$ where $\mu = w_1\mu_1 + ... + w_J\mu_J$ and $\sigma^2 = w_1^2\sigma_1^2 + ... + w_J^2\sigma_J^2$
- Sample Mean and Variance

- $\bar{X} = \frac{1}{N}\sum_{n=1}^{N} X_n$
- $s_X^2 = \frac{1}{N-1}\sum_{n=1}^{N}(X_n - \bar{X})^2$
- $s_X^2 = \frac{N}{N-1}(\overline{X^2} - \bar{X}^2)$

- X and Y are independent if,
  - $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ (discrete case)
  - $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ (continuous case)

- Covariance and Correlation
  - $\sigma_{XY} = Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - E[X]E[Y]$
  - $\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

- Linear Combinations
  - $Cov(aX+b, Y) = aCov(X,Y)$
  - $Cov(X, aY+b) = aCov(X,Y)$

- Conditional Distributions
  - $p_{Y|X}(y|x) = \frac{p_{X,Y}(x,y)}{p_X(x)}$ (discrete case)
  - $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$ (continuous case)

- Conditional Expectations
  - $E[Y|X] = \sum_y y p_{Y|X}(y|x) = \frac{\sum_y y p_{X,Y}(x,y)}{\sum_y p_{X,Y}(x,y)}$
  - $E[Y|X] = \int_y y f_{Y|X}(y|x)dy = \frac{\int_y y f_{X,Y}(x,y)dy}{\int_y f_{X,Y}(x,y)dy}$
  - $E[X] = E_Y[E[X|Y]]$ (Law of Iterated Expectations)

- Bivariate Normal Distribution:

$$f_{X,Y}(x,y;\mu_X,\sigma_X^2,\mu_Y,\sigma_Y^2,\rho)$$

$$= \frac{1}{\sigma_X\sigma_Y 2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}}$$

- Sampling from the Normal Distribution: If $X_n \sim N(\mu_X, \sigma_X^2)$ and X_n are independent, then,
  - $E[\bar{X}] = \mu_X$
  - $Var(\bar{X}) = \frac{1}{N}\sigma_X^2$
  - $\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{N}} \sim N(0,1)$
  - $\frac{(N-1)s_X^2}{\sigma_X^2} \sim \chi^2_{N-1}$
  - $\bar{X}$ and s_X^2 are independent
- The Derived Distributions
  - If $U \sim N(0,1)$, $V \sim \chi^2_v$, and U and V are independent, then $\frac{U}{\sqrt{V/v}} \sim t_v$
  - If $U \sim \chi^2_u$, $V \sim \chi^2_v$, and U and V are independent, then $\frac{U/u}{V/v} \sim F_{u,v}$
- Sampling Distributions
  - Mean with known variance, $Z = \frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{N}} \sim N(0,1)$
  - Mean with unknown variance, $t = \frac{\bar{X}-\mu_X}{s_X/\sqrt{N}} \sim t_{N-1}$
- Asymptotic Theory
  - $\xrightarrow{prob.}$ = convergence in probability
  - $\xrightarrow{dist.}$ = convergence in distribution
  - $\bar{X} \xrightarrow{prob.} E[X]$ (LLN)
  - $\frac{1}{N}\sum_{n=1}^{N} g(X_n) \xrightarrow{prob.} E[g(X)]$ (LLN)
  - $\frac{\sqrt{N}(\bar{X}-\mu_X)}{\sigma_X} \xrightarrow{dist.} N(0,1)$ (CLT)
  - $\frac{1}{\sqrt{N}}\sum_{n=1}^{N}(g(X_n) - E[g(X_n)]) \xrightarrow{dist.} N(0, Var(g(X_n))$ (CLT)
- Large Sample Sampling Distributions
  - $Z = \frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{N}} \xrightarrow{dist.} N(0,1)$
  - $t = \frac{\bar{X}-\mu_X}{s_X/\sqrt{N}} \xrightarrow{dist.} N(0,1)$

- Vector Probability Theory
  - $\mu = E[X]$
  - $\Omega = Var(X) = E[(X-\mu)(X-\mu)']$
  - $Var(X) = E[(X-\mu)(X-\mu)'] = E[XX'] - E[X]E[X]'$
  - $E[AX+b] = AE[X] + b$
  - $Var(AX+b) = AVar(X)A'$
  - $\bar{X} \xrightarrow{prob.} E[X]$ (LLN)
  - $\frac{1}{N}\sum_{n=1}^{N} g(X_n) \xrightarrow{prob.} E[g(X)]$ (LLN)
  - $\sqrt{N}(\bar{X}-\mu) \xrightarrow{dist.} N(0,\Omega)$ (CLT)
  - $\frac{1}{\sqrt{N}}\sum_{n=1}^{N}(g(X_n) - E[g(X_n)]) \xrightarrow{dist.} N(0, Var(g(X_n))$ (CLT)
  - Multivariate normal distribution:

$$f_X(x;\mu,\Omega) = \frac{1}{(2\pi)^{K/2}(\det\Omega)^{1/2}} e^{-\frac{1}{2}(x-\mu)'\Omega^{-1}(x-\mu)}$$