Mark's Formula Sheet for Exam P

Discrete distributions

- Uniform, $U(m)$
  - PMF: $f(x) = \frac{1}{m}$, for $x = 1, 2, \ldots, m$
  - $\mu = \dfrac{m+1}{2}$ and $\sigma^2 = \dfrac{m^2-1}{12}$
- Hypergeometric
  - PMF: $f(x) = \dfrac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$
  - x is the number of items from the sample of n items that are from group/type 1.
  - $\mu = n(\dfrac{N_1}{N})$ and $\sigma^2 = n(\dfrac{N_1}{N})(\dfrac{N_2}{N})(\dfrac{N-n}{N-1})$
- Binomial, $b(n,p)$
  - PMF: $f(x) = \dbinom{n}{x} p^x (1-p)^{n-x}$, for $x = 0, 1, \ldots, n$
  - x is the number of successes in n trials.
  - $\mu = np$ and $\sigma^2 = np(1-p) = npq$
  - MGF: $M(t) = [(1-p) + pe^t]^n = (q + pe^t)^n$
- Negative Binomial, $nb(r,p)$
  - PMF: $f(x) = \dbinom{x-1}{r-1} p^r (1-p)^{x-r}$, for $x = r, r+1, r+2, \ldots$
  - x is the number of trials necessary to see r successes.
  - $\mu = r(\dfrac{1}{p}) = \dfrac{r}{p}$ and $\sigma^2 = \dfrac{r(1-p)}{p^2} = \dfrac{rq}{p^2}$
  - MGF: $M(t) = \dfrac{(pe^t)^r}{[1-(1-p)e^t]^r} = \left(\dfrac{pe^t}{1-qe^t}\right)^r$
- Geometric, $geo(p)$
  - PMF: $f(x) = (1-p)^{x-1}p$, for $x = 1, 2, \ldots$
  - x is the number of trials necessary to see 1 success.
  - CDF: $P(X \le k) = 1 - (1-p)^k = 1 - q^k$ and $P(X > k) = (1-p)^k = q^k$
  - $\mu = \dfrac{1}{p}$ and $\sigma^2 = \dfrac{1-p}{p^2} = \dfrac{q}{p^2}$
  - MGF: $M(t) = \dfrac{pe^t}{1-(1-p)e^t} = \dfrac{pe^t}{1-qe^t}$
  - Distribution is said to be "memoryless", because $P(X > k+j | X > k) = P(X > j)$.

- Poisson
  - PMF: $f(x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$, for $x = 0, 1, 2, \ldots$
  - x is the number of changes in a unit of time or length.
  - λ is the average number of changes in a unit of time or length in a Poisson process.
  - CDF: $P(X \leq x) = e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^x}{x!})$
  - $\mu = \sigma^2 = \lambda$
  - MGF: $M(t) = e^{\lambda(e^t - 1)}$

Continuous Distributions

- Uniform, $U(a, b)$
  - PDF: $f(x) = \dfrac{1}{b-a}$, for $a \leq x \leq b$
  - CDF: $P(X \leq x) = \dfrac{x-a}{b-a}$, for $a \leq x \leq b$
  - $\mu = \dfrac{a+b}{2}$ and $\sigma^2 = \dfrac{(b-a)^2}{12}$
  - MGF: $M(t) = \dfrac{e^{tb} - e^{ta}}{t(b-a)}$, for $t \neq 0$, and $M(0) = 1$
- Exponential
  - PDF: $f(x) = \dfrac{1}{\theta} e^{-x/\theta}$, for $x \geq 0$
  - x is the waiting time we are experiencing to see one change occur.
  - θ is the average waiting time between changes in a Poisson process. (Sometimes called the "hazard rate".)
  - CDF: $P(X \leq x) = 1 - e^{-x/\theta}$, for $x \geq 0$.
  - $\mu = \theta$ and $\sigma^2 = \theta^2$
  - MGF: $M(t) = \dfrac{1}{1 - \theta t}$
  - Distribution is said to be "memoryless", because $P(X \geq x_1 + x_2 | X \geq x_1) = P(X \geq x_2)$.
- Gamma
  - PDF: $f(x) = \dfrac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta} = \dfrac{1}{(\alpha-1)!\theta^\alpha} x^{\alpha-1} e^{-x/\theta}$, for $x \geq 0$
  - x is the waiting time we are experiencing to see α changes.
  - θ is the average waiting time between changes in a Poisson process and α is the number of changes that we are waiting to see.
  - $\mu = \alpha\theta$ and $\sigma^2 = \alpha\theta^2$
  - MGF: $M(t) = \dfrac{1}{(1 - \theta t)^\alpha}$
- Chi-square (Gamma with $\theta = 2$ and $\alpha = \frac{r}{2}$)
  - PDF: $f(x) = \dfrac{1}{\Gamma(r/2) 2^{r/2}} x^{r/2-1} e^{-x/2}$, for $x \geq 0$
  - $\mu = r$ and $\sigma^2 = 2r$
  - MGF: $M(t) = \dfrac{1}{(1 - 2t)^{r/2}}$

- Normal, $N(\mu, \sigma^2)$
  - PDF: $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$
  - MGF: $M(t) = e^{\mu t + \sigma^2 t^2/2}$

Integration formulas

- $\displaystyle\int p(x)e^{ax}\ dx = \frac{1}{a}p(x)e^{ax} - \frac{1}{a^2}p'(x)e^{ax} + \frac{1}{a^3}p''(x)e^{ax} - \ldots$
- $\displaystyle\int_a^\infty x\left(\frac{1}{\theta}e^{-x/\theta}\right)\ dx = (a+\theta)e^{-a/\theta}$
- $\displaystyle\int_a^\infty x^2\left(\frac{1}{\theta}e^{-x/\theta}\right)\ dx = ((a+\theta)^2+\theta^2)e^{-a/\theta}$

Other Useful Facts

- $\sigma^2 = E[(X-\mu)^2] = E[X^2] - \mu^2 = M''(0) - M'(0)^2$
- $\mathrm{Cov}(X,Y) = E[(X-\mu_x)(Y-\mu_y)] = E[XY] - \mu_x\mu_y$
- $\mathrm{Cov}(X,Y) = \sigma_{xy} = \rho\sigma_x\sigma_y$
  and
  $\rho = \dfrac{\sigma_{xy}}{\sigma_x\sigma_y}$
- Least squares regression line: $y = \mu_y + \rho\dfrac{\sigma_y}{\sigma_x}(x-\mu_x)$
- When variables $X_1, X_2, \ldots, X_n$ are not pairwise independent, then
  $\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \sigma_i^2 + 2\sum_{i<j}\sigma_{ij}$
  and
  $\mathrm{Var}(\sum_{i=1}^n a_iX_i) = \sum_{i=1}^n a_i^2\sigma_i^2 + 2\sum_{i<j}a_ia_j\sigma_{ij}$
  where σ_{ij} is the covariance of X_i and X_j.
- When X depends upon Y, $E[X] = E[E[X|Y]]$.
- When X depends upon Y, $\mathrm{Var}(X) = E[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y])$. (Called the "Total Variance" of X.)
- Chebyshev's Inequality: For a random variable X having any distribution with finite mean μ and variance σ^2, $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$.
- For the variables X and Y having the joint PMF/PDF $f(x,y)$, the moment generating function for this distribution is

  $$M(t_1,t_2) = E[e^{t_1X+t_2Y}] = E[e^{t_1X}e^{t_2Y}] = \sum_x\sum_y e^{t_1x}e^{t_2y}f(x,y)$$

  - $\mu_x = M_{t_1}(0,0)$ and $\mu_y = M_{t_2}(0,0)$ (These are the first partial derivatives.)
  - $E[X^2] = M_{t_1t_1}(0,0)$ and $E[Y^2] = M_{t_2t_2}(0,0)$ (These are the "pure" second partial derivatives.)
  - $E[XY] = M_{t_1t_2}(0,0) = M_{t_2t_1}(0,0)$ (These are the "mixed" second partial derivatives.)
- Central Limit Theorem: As the sample size n grows,

- the distribution of $\sum_{i=1}^{n} X_i$ becomes approximately normal with mean $n\mu$ and variance $n\sigma^2$
- the distribution of $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ becomes approximately normal with mean μ and variance $\frac{\sigma^2}{n}$.

- If X and Y are joint distributed with PMF $f(x, y)$, then
  - the marginal distribution of X is given by $f_x(x) = \sum_y f(x, y)$
  - the marginal distribution of Y is given by $f_y(y) = \sum_x f(x, y)$
  - $f(x|y = y_0) = \frac{f(x, y_0)}{f_y(y_0)}$.
  - $E[X|Y = y_0] = \sum_x xf(x|y = y_0) = \frac{\sum_x xf(x, y_0)}{f_y(y_0)}$