STAT 330: Formula Sheet for the Final Exam

- If $A, B \in \mathscr{B}$, then $P(A^c) = 1 - P(A)$ and $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Markov: $P(|X| \geq c) \leq \frac{E(|X|^k)}{c^k}, \forall k > 0, \forall c > 0$, Chebyshev: $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}, \forall k > 0$
- $E(X) = E\Big(E(X|Y)\Big)$, $Var(X) = E\Big(Var(X|Y)\Big) + Var\Big(E(X|Y)\Big)$
- Defining $U = h_1(X,Y)$ and $V = h_2(X,Y)$ where $X = w_1(U,V)$ and $Y = w_2(U,V)$, then
$$f_{U,V}(u,v) = f_{X,Y}\Big(w_1(u,v)\,,\, w_2(u,v)\Big) \times |J| \;,\quad (u,v) \in Support(U,V) \quad \text{where } J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$
- If $(X_1, X_2) \sim BVN(\mu, \Sigma)$ then $X_2|X_1 = x_1 \sim \mathrm{N}\Big(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)\,,\, \sigma_2^2(1-\rho^2)\Big)$
- δ-method: If $n^b(X_n - a) \xrightarrow{\mathrm{D}} X$, then $n^b\big[g(X_n) - g(a)\big] \xrightarrow{\mathrm{D}} g'(a)X$, where $g(x)$ is differentiable at a, $g'(a) \neq 0$, and $b > 0$.
- If $X_n \xrightarrow{\mathrm{P}} a$ and $g(x)$ is continuous at $x = a$, then $g(X_n) \xrightarrow{\mathrm{P}} g(a)$.
- If $X_n \xrightarrow{\mathrm{P}} a$, $Y_n \xrightarrow{\mathrm{P}} b$ and $g(x,y)$ is continuous at (a,b) then $g(X_n, Y_n) \xrightarrow{\mathrm{P}} g(a,b)$.
- (Continuous Mapping/Slutsky's Theorem) If $X_n \xrightarrow{\mathrm{D}} X$, $Y_n \xrightarrow{\mathrm{P}} b$, and $g(x,b)$ is continuous for all x in the support set of X, then $g(X_n, Y_n) \xrightarrow{\mathrm{D}} g(X,b)$.
- If $X_n \xrightarrow{\mathrm{D}} X$ and $g(x)$ is continuous for all $x \in$ support set of X, then $g(X_n) \xrightarrow{\mathrm{D}} g(X)$.
- For a scaler parameter θ, we have $S(\theta) = \frac{d\log(L(\theta))}{d\theta}$, $I(\theta) = \frac{-d^2\log(L(\theta))}{d\theta^2}$, $J(\theta) = E(I(\theta; X))$

- $\displaystyle\sum_{x=0}^{\infty} a^x = \frac{1}{1-a}$ if $|a| < 1$ and $\displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$
- $\Gamma(z) = \displaystyle\int_0^{\infty} e^{-t}t^{z-1}dt = (z-1)!$
- $(fg)' = f'g + fg'$ and $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$
- $(f(g(x)))' = f'(g(x))g'(x)$
- $\displaystyle\int_a^b u\,dv = [uv]_a^b - \int_a^b v\,du$
- For non-neg X, $\mathbb{E}[X] = \displaystyle\int_0^{\infty} P(X \geq x)dx \overset{or}{=} \sum_{x=1}^{\infty} P(X \geq x)$
- $M_{aX+b}(t) = e^{bt}M_X(at)$
- $\mathrm{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
- $\mathrm{Var}(aX + bY + c) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X,Y)$
- $-1 \leq \rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \leq 1$
- If $\mathbf{X} \sim BVN(\mu, \Sigma)$ then $A\mathbf{X} + \mathbf{b} \sim BVN(A\mu + \mathbf{b}, A\sigma A^T)$
- If $(X_1, \ldots, X_k) \sim Mult(n, p_1, \ldots, p_k)$ then $X_i|X_j = x_j \sim Bin(n - x_j, \frac{p_i}{1-p_j})$ and $X_i|X_i + X_j = t \sim Bin(t, \frac{p_i}{p_i+p_j})$
- If $\lim_{n\to\infty}\psi(n) = 0$ then $\lim_{n\to\infty}(1 + \frac{b}{n} + \frac{\psi(n)}{n})^{cn} = e^{bc}$.
- If $X_1, X_2, \ldots$ is an i.i.d. sequence with $\mathbb{E}[X_1] = \mu < \infty$ and $\mathrm{Var}(X_1) = \sigma^2 < \infty$. Let $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$. (WLLN) Then $\lim_{n\to\infty}\overline{X_n} \xrightarrow{P} \mu$. (CLT) Then $\frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$.
- Likelihood and score: $L(\theta_1, \ldots, \theta_k; x) = \prod_{i=1}^n f_X(x_i; \theta_1, \ldots, \theta_k)$ and $\ell(\theta_1, \ldots, \theta_k; x) = \sum_{i=1}^n \log f_X(x_i; \theta_1, \ldots, \theta_k)$ and $S(\theta_1, \ldots, \theta_k; x) = \left[\frac{\partial\ell}{\partial\theta_1}, \ldots, \frac{\partial\ell}{\partial\theta}\right]$.
- Information: $k \times k$ symmetric matrix $I(\theta)_{i,j} = -\frac{\partial^2\ell(\theta_1,\ldots,\theta_k)}{\partial\theta_i\partial\theta_j}$, expected information is $J(\theta)_{i,j} = \mathbb{E}[I(\theta)_{i,j}]$. Check PSD by $\det(I(\hat{\theta})) > 0$.
- $R(\theta; x) = \frac{L(\theta)}{L(\hat{\theta})}$, $100p\%$ likelihood region is $\{\theta : R(\theta) \geq p\}$.
- If $p = 2(Z \leq a) - 1$ then $\{\theta : R(\theta) \geq e^{-a^2/2}\}$ is a $100p\%$ CI for θ.
- Pivotal quantities: (1) $\bar{X}_n - \mu \sim N(0, \frac{\sigma^2}{n})$, (2) $\frac{\bar{X}_n - \mu}{s/\sqrt{n}} \sim t_{(n-1)}$, (3) $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$, (4) if θ is location (resp. scale) family then $Q = \tilde{\theta} - \theta$ (resp. $Q = \tilde{\theta}/\theta$) is pivotal.
- $\Lambda(X) = -2\log\frac{\max_{\theta\in\Omega_0} L(\theta;X)}{\max_{\theta\in\Omega} L(\theta;X)} = 2(\ell(\tilde{\theta};X) - \max_{\theta\in\Omega_0}\ell(\theta;X))$.

- **σ-Field:** $\mathcal{F} \subseteq \mathfrak{P}(\Omega)$ such that (1) $\emptyset, \Omega \in \mathcal{F}$, (2) if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$, (3) if $A_1, A_2, \ldots \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.
- **Random Variable:** $X : \Omega \to \mathbb{R}$ is an R.V. if for all $x \in \mathbb{R}$, $P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\})$ is defined.
- For X continuous, $F(X) \sim U(0,1)$ and $F^{-1}(U) \sim X$.
- L-S family: $F_X(x) = F_0((x-\mu)/\sigma)$ or $f_X(x) = \frac{1}{\sigma}f_1((x-\mu)/\sigma)$.
- Marginalizing: $f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\infty}^{\infty} f(x|y)f_Y(y)dy$ and $F_X(x) = \lim_{y\to\infty} F(x,y)$.
- The following are equivalent:
  - X and Y are independent.
  - $F(x,y) = F_X(x)F_Y(y)$.
  - $f(x,y) = f_X(x)f_Y(y)$.
  - $f(x|y) = f_X(x)$ or $f(y|x) = f(y)$.
  - $M(t_1,t_2) = M_X(t_1)M_Y(t_2)$.
  - $supp(X,Y) = supp(X) \times supp(Y)$ and there are functions h, g such that $f(x,y) = h(x)g(y)$.

  Note that $supp(X,Y) = supp(X) \times supp(Y)$ is necessary but not sufficient.
- $X \perp Y$ then $\mathbb{E}[g(X)g(Y)] = \mathbb{E}[g(X)]\mathbb{E}[g(Y)] \implies \mathrm{Cov}(X,Y) = 0$.
- If $X_1, \ldots, X_n$ ind. and $Y = \sum_{i=1}^n X_i$ then $M_Y(t) = \prod_{i=1}^n M_{X_i}(t)$
- $X_n \xrightarrow{d} X$ iff $\lim_{n\to\infty} F_n(x) = F_X(x)$ iff $\lim_{n\to\infty} M_n(t) = M_X(t)$ and $X_n \xrightarrow{P} X$ if $\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$.
- If f is one-to-one then $\mathrm{MLE}(f(\theta)) = f(\mathrm{MLE}(\theta))$.
- Regularity conditions (when support doesn't depend on θ):
  - Consistency: $\tilde{\theta}_n \xrightarrow{P} \theta$.
  - Asymptotic normality: $\sqrt{J(\theta)}(\tilde{\theta}_n - \theta) \xrightarrow{D} N(0, I_k)$ (also holds for any of $J(\theta)$, $J(\tilde{\theta})$, $I(\theta)$, or $I(\tilde{\theta})$). Note then $\mathrm{Var}(\tilde{\theta}) \approx J(\hat{\theta})^{-1}$ for large n.
  - RLL: $-2\log R(\theta) = 2(\ell(\tilde{\theta}) - \ell(\theta)) \xrightarrow{D} \chi^2_{(k)}$.
  - $J(\theta) = \mathbb{E}[-\frac{\partial^2\ell(\theta)}{\partial\theta^2}] = \mathbb{E}[(\frac{\partial\ell(\theta)}{\partial\theta})^2] = \mathrm{Var}(S(\tilde{\theta}))$.
- Confidence intervals: Find a pivotal quantity to get $P(q_1 \leq Q(X;\theta) \leq q_2) = 1 - \alpha$ and then $P(A(X) \leq \theta \leq B(X)) = 1 - \alpha$.
- Likelihood ratio test: (1) set up $H_0 : \theta \in \Omega_0$ and H_a, (2) find $\tilde{\theta}_{ML}$, (3) find $\lambda(\theta_0)$, (4) get DoF = # of parameters - # of constraints, (5) get p-value, (6) interpret.