Stats Final Exam Cheat Sheet

Gamma function

$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$; for $\mathcal{R}(z) > 0$
$\Gamma(z+1) = z\Gamma(z)$; $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$

Gamma distribution

$f_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1}e^{-\beta y}$; $y, \alpha, \beta > 0$
$E[Y] = \alpha/\beta$; $Var(Y) = \alpha/\beta^2$
$\text{MGF}(t) = \left(1 - \frac{t}{\beta}\right)^{-\alpha}$ for $t < \beta$
$\chi^2_k \sim \text{Gamma}(k/2, 1/2)$
Y_i indip. $\text{Gamma}(\alpha_i, \beta) \Rightarrow \sum_{i=1}^n Y_i \sim \text{Gamma}(\sum_{i=1}^n \alpha_i, \beta)$
$k * \text{Gamma}(\alpha, \beta) = \text{Gamma}(\alpha, \beta/k)$

Inverse Gamma distribution

$X \sim \text{Gamma}(\alpha, \beta)$ $Y = 1/X \sim \text{I-Gamma}(\alpha, \beta)$
$f_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1}e^{-\beta/y}$; $y, \alpha, \beta > 0$
$E[Y] = \beta/(\alpha-1)$; $Var(Y) = \beta^2/[(\alpha-1)^2(\alpha-2)]$
Mode $= \beta/(\alpha+1)$

Inverse Gaussian distribution

$Y \sim \text{I-G}(\phi, \lambda)$
$f_Y(y) = \frac{\sqrt{\lambda}}{\sqrt{2\pi}} e^{\sqrt{\lambda\phi}} y^{-3/2} \exp\{-\frac{1}{2}\left(\phi y + \frac{\lambda}{y}\right)\}$;
$y, \phi \geq 0, \lambda > 0$
$E[Y] = \left(\frac{\lambda}{\phi}\right)^{1/2}$; $Var(Y) = \sqrt{\frac{\lambda}{\phi^3}}$

Normal Inverse Gamma distribution

$\theta \sim \text{N-IGa}(\eta_0, \lambda_0, \alpha_0, \beta_0)$
$\sigma^2 \sim \text{IGa}(\alpha_0, \beta_0)$
$\mu|\sigma^2 \sim \text{N}(\eta_0, \sigma^2/\lambda_0)$
$f(\mu, \sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha_0+3/2} \exp\{\lambda_0\eta_0\frac{\mu}{\sigma^2} - \frac{1}{2\sigma^2}(\lambda_0\eta_0^2 + 2\beta_0) - \lambda_0\frac{\mu^2}{2\sigma^2}\}$
Posterior params (from Normal model s.r.s.):
$\alpha = \alpha_0 + \frac{n}{2}$, $\lambda = \lambda_0 + n$
$\eta = \frac{\lambda_0\eta_0 + n\bar{y}}{\lambda_0 + n}$
$\beta = \beta_0 + \frac{n}{2}\upsilon^2 + \frac{n\lambda_0}{2(\lambda_0+n)}(\bar{y} - \eta_0)^2$
$\upsilon = n^{-1}\sum_{i=1}^n (y_i - \bar{y})^2$

Beta distribution

$f_Y(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}$; $\alpha, \beta > 0$, $y \in [0, 1]$
$E[Y] = \frac{\alpha}{\alpha+\beta}$; $Var(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Multivariate Normal distribution

$f_Y(y)\frac{\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)}{2\pi^{d/2}\det(\Sigma)^{1/2}}$; $\text{MGF}(t) = \exp\{\mu^T t + \frac{1}{2}t^T\Sigma t\}$; $Y_1|Y_2 = y_2 \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y_2 - \mu_2); \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$
$X|\mu \sim N(\mu, \sigma^2)$ and $\mu \sim N(\gamma, \tau^2)$
$X \sim N(\gamma, \sigma^2 + \tau^2)$; $y|\theta \sim N(\theta, \sigma_0^2)$ (σ_0^2 known) $\theta \sim N(\mu_0, \tau_0^2) \Rightarrow \theta|y \sim N(\mu, \tau^2)$
$\frac{1}{\tau^2} = \frac{1}{\tau_0^2} + \frac{n}{\sigma_0^2}$; $\mu = \mu_0\frac{\tau^2}{\tau_0^2} + \bar{y}\frac{n\tau^2}{\sigma_0^2}$

t-Student distribution

$f_Y(y) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2}$
$\nu > 0$

Uniform distribution

$Y \sim Unif(a, b)$
$f_Y(y) = \frac{1}{b-a}$ for $y \in [a, b]$ otherwise 0
$F_Y(y) = \frac{y-a}{b-a}$ for $y \in [a, b]$
$E[Y] = \frac{1}{2}(a+b)$; $Var(Y) = \frac{1}{12}(b-a)^2$

Cauchy distribution

$Y \sim Cau(\mu, \sigma)$
$f_Y(y) = \left\{\pi\sigma\left(1 + \left(\frac{y-\mu}{\sigma}\right)^2\right)\right\}^{-1}$

F distribution

$Y \sim F(\nu_1, \nu_2)$ $y, \nu_1, \nu_2 > 0$
$f_Y(y) = \frac{(\nu_1/\nu_2)^{\nu_1/2}y^{\nu_1/2-1}}{B(\nu_1/2, \nu_2/2)[1+(\nu_1/\nu_2)y]^{(\nu_1+\nu_2)/2}}$
$E[Y] = \frac{\nu_2}{\nu_2 - 2}$

Logistic distribution

$Y \sim L(\mu, \sigma)$; $\mu \in \mathbb{R}$ $f(y) = \frac{e^{-(y-\mu)/\sigma}}{\sigma(1+e^{-(y-\mu)/\sigma})^2}$; $E[Y] = \text{Median} = \text{Mode} = \mu$; $Var(Y) = \frac{\sigma^2\pi}{3}$

Log Normal distribution

$Y \sim LN(\mu, \sigma)$; $y > 0$
$f(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log(y)-\mu)^2}{2\sigma^2}\right\}$
$E[Y] = \exp\left\{\mu + \frac{\sigma^2}{2}\right\}$; Mode $= exp\left\{\mu - \sigma^2\right\}$

Hypergeometric distribution

$f_Y(y) = \frac{\binom{K}{y}\binom{N-K}{n-y}}{\binom{N}{n}}$
$E[Y] = n\frac{K}{N}$; $Var(Y) = n\frac{N-K}{N}\frac{K}{N}\frac{N-n}{N-1}$

Negative Binomial distribution

$Y =$ "n. of failures till r-th success"
$f_Y(k) = \binom{k+r-1}{k}(1-p)^k p^r$; $\frac{\Gamma(k+r)}{k!\Gamma(r)}$
$E[Y] = \frac{r(1-p)}{p}$; $Var(Y) = \frac{r(1-p)}{p^2}$

Weibull distribution

$Y \sim Wei(\gamma, \beta)$ $y, \beta, \gamma > 0$
$f(y) = \gamma\beta^{-\gamma}y^{\gamma-1}e^{-(y/\beta)^\gamma}$
$E[Y] = \beta\Gamma(1 + 1/\gamma)$

Series

$\sum_{i=0}^n r^i = \frac{1-r^{n+1}}{1-r}$ for $|r| < 1$
$e^x = \sum_{n=0}^\infty x^n/n! = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$
$\sum_{i=0}^n \frac{1}{x^p}$ for $p > 1$
$\int_0^\infty y^\alpha e^{-ky}dy < \infty$ if $\alpha > -1$ ($k > 0$)

Log base change

$\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$

Taylor expansions

$e^x = \sum_{i=0}^k \frac{x^i}{i!} + o(x^k)$ for $x \to 0$
$\log(1+x) = \sum_{i=1}^k (-1)^{i-1}\frac{x^i}{i} + o(x^k)$ for $x \to 0$

Matrix

$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \xrightarrow{Inverse} \frac{1}{\det X}\begin{pmatrix} D & -B \\ -C & A \end{pmatrix}$
$X_{2\times 2}$ is positive definite if $A > 0$ and $\det X > 0$

Reparam and Block

Reparam: $U^\Gamma(\gamma) = D(\gamma)^T U^\Theta(\theta(\gamma))$
$j^\Gamma(\hat{\gamma}) = D(\hat{\gamma})^T j^\Theta(\hat{\theta})D(\hat{\gamma})$
$i^\Gamma(\gamma) = D(\gamma)^T i^\Theta(\theta(\gamma))D(\gamma)$
$D(\gamma) = \left[\frac{\delta\theta_r(\gamma)}{\delta\gamma_s}\right]$
Block: $i^{\psi\psi} = (i_{\psi\psi} - i_{\psi\lambda}(i_{\lambda\lambda})^{-1}i_{\lambda\psi})^{-1}$
$i^{\psi\lambda} = -i^{\psi\psi}i_{\psi\lambda}i_{\lambda\lambda}^{-1}$

Order Statistics

i.i.d continuous case.
$f_{Y_{(1)},..,Y_{(n)}}(y_1, ., y_n) = n!f(y_1)f(y_2)...f(y_n)$
$f_{Y(r)}(y_r) = \frac{n![F(y_r)]^{r-1}f(y_r)[1-F(y_r)]^{n-r}}{(n-r)!(r-1)!}$
$f_{Y(r)Y(s)}(y, z) = \frac{n![F(y)]^{r-1}f(y)[F(z)-F(s)]^{s-r-1}f(z)[1-F(z)]^{n-r}}{(n-s)!(s-r-1)!(r-1)!}$
$F(y_{(r)})]^{n-r}$
From n choose first r $f_{Y_{(1)},..,Y_{(r)}}(y_1, ., y_r) = \frac{n!}{(n-r)!}([1 - F(y_{(r)})]^{n-r})\prod_{i=1}^r f(y_{(i)})$

Profile pivot

Likelihood profile pivot $\theta = (\psi, \lambda)$
$W_{P_e}(\psi) = (\hat{\psi} - \psi)^T j_P(\hat{\psi})(\hat{\psi} - \psi)$
$W_{P_u}(\psi) = U_P(\psi)^T i_P^{\psi\psi}(\psi, \hat{\lambda}_\psi)U_P(\psi)$
we can use $j^{\psi\psi}$ (see reparam and block)

Confidence intervals and Test

Binomial: for $n\theta$, $n(1-\theta) > 5$
$q_2(\theta, \hat{\theta}) = \frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}}$
$\frac{\hat{\theta} + (z_{1-\alpha/2})^2/2n}{1 + (z_{1-\alpha/2})^2/n} \pm \frac{z_{1-\alpha/2}/\sqrt{n}}{1 + (z_{1-\alpha/2})^2/n}\sqrt{\hat{\theta}(1-\hat{\theta}) + \frac{(z_{1-\alpha/2})^2}{4n}}$
Clopper Pearson:
$\{\theta \in (0, 1) : \mathbb{P}_\theta(Y \leq y) \geq \alpha/2$ and $\mathbb{P}_\theta(Y \geq y) \geq \alpha/2\}$
Exact distribution test: $n\hat{\theta} \sim Bi(n, \theta)$
Mid-pvalue: $\alpha_{mid}^{oss} = \sup_{\theta\in\Theta_0}\{\mathbb{P}_\theta(t(Y) > t(y)) + \frac{1}{2}\mathbb{P}_\theta(t(Y) = t(y))\}$

Netwon - Raphson

$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
$\hat{\theta}_{s+1} = \hat{\theta}_s + j(\hat{\theta}_s)^{-1}U(\hat{\theta}_s)$

Total Variance

$Var(Y) = E[Var(Y|X)] + Var(E[Y|X])$

Delta Method

$\sqrt{n}(Y_n - \theta) \xrightarrow{D} N_d(0, \Sigma)$
$\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{D} N_k(0, D\Sigma D^T)$
$D_{k \times d} = [d_{ij}]$ $d_{ij} = \frac{\delta g_i(\theta)}{\delta \theta_j}$

Posterior Predicitive Density

if $(y^*|y, \theta) = (y^*|\theta)$
$p(y^*|y) = \int p(y^*|\theta) p(\theta|y) d\theta$

Another topic

some stuff