MATH 324 Statistics - Midterm Crib Sheet

Expectation: $\mathbb{E}[g(Y)] = \sum_y g(y)p(y)$ discrete, $\int_{-\infty}^{\infty} g(y)f(y)\mathrm{d}y$ continuous; $\mathbb{E}[\overline{Y}^2] = \mu^2 + \frac{\sigma^2}{n}$
Variance:
$\mathbb{V}[Y] = \mathbb{E}[(Y - \mathbb{E}[Y])^2] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$, $\mathbb{V}[c] = 0$, $\mathbb{V}[Y + c] = \mathbb{V}[Y]$, $\mathbb{V}[cY] = c^2\mathbb{V}[Y]$, $\mathbb{V}[X + Y] = \mathbb{V}[X] + 2\mathrm{Cov}[X, Y] + \mathbb{V}[Y] = \mathbb{V}[X] + \mathbb{V}[Y]$ if X and Y indep, $\mathbb{V}[aX + bY + c] = a^2\mathbb{V}[X] + 2ab\mathrm{Cov}[X, Y] + b^2\mathbb{V}[Y]$
$\mathbb{V}\left[\overline{Y}\right] = 1/n^2\mathbb{V}[\sum Y_i] = 1/n^2\left(\sum \mathbb{V}[Y_i] + 2\sum\sum_{1\leq i\leq j\leq n} \mathrm{Cov}(Y_i, Y_j)\right) = 1/n^2\sum \mathbb{V}[Y_i] = n\sigma^2/n^2 = \sigma^2/n$
Covariance: $\mathrm{Cov}[Y_1, Y_2] = \mathbb{E}[(Y_1 - \mathbb{E}[Y_1])(Y_2 - \mathbb{E}[Y_2])] = \mathbb{E}[Y_1Y_2] - \mathbb{E}[Y_1]\mathbb{E}[Y_2]$, $\mathrm{Cov}[aY_1, bY_2] = ab\mathrm{Cov}[Y_1, Y_2]$. $Z \sim N(0, 1) \Rightarrow \mathrm{Cov}[Z, Z^2] = 0$
Correlation: $\mathrm{Corr}[Y_1, Y_2] = \frac{\mathrm{Cov}[Y_1,Y_2]}{\sqrt{\mathbb{V}[Y_1]\mathbb{V}[Y_2]}}$. Joint pdf symmetric on Y_1 and $Y_2 \Rightarrow$ same marginal distribs & $\mathbb{E} \Rightarrow$ no correlation.

Standard deviation: $\sigma = \sqrt{\mathbb{V}[Y]}$. **Sample std deviation:** $s = \sqrt{s^2} = \sqrt{\sum(Y - \overline{Y})^2/(n-1)}$; sample variance s^2 unbiased estimator of σ^2
MSE($\hat{\theta}$)**:** $\mathbb{E}[(\hat{\theta} - \theta)^2] = \mathbb{E}[\hat{\theta}^2] - 2\theta\mathbb{E}[\hat{\theta}] + \theta^2$; $\mathbb{V}\left[\hat{\theta}\right] = \mathrm{MSE}[\hat{\theta}]$ if $\hat{\theta}$ unbiased estimator of θ
Expected vals and std errors of some common point estimators:

target param θ	sample size(s)	point estimator $\hat{\theta}$	$\mathbb{E}[\hat{\theta}]$	std error $\sigma_{\hat{\theta}}$
μ	n	$\overline{Y}$	μ	$\frac{\sigma}{\sqrt{n}}$
p	n	$\hat{p} = Y/n$	p	$\sqrt{\frac{pq}{n}}$
$\mu_1 - \mu_2$	n_1 and n_2	$\overline{Y}_1 - \overline{Y}_2$	$\mu_1 - \mu_2$	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$p_1 - p_2$	n_1 and n_2	$\hat{p}_1 - \hat{p}_2$	$p_1 - p_2$	$\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$

2-standard-error: $2\sigma_{\hat{\theta}}$
CI ($100(1-\alpha)\%$ **confidence interval):**
$P(a \leq U \leq b) = 1-\alpha$, $P(U \leq a) = \int_0^a f_U(u)\mathrm{d}u = P(U \geq b) = \int_b^{\infty} f_U(u)\mathrm{d}u = \alpha/2$. In **large samples**, estimators have **normal sampling distributions:** $\mathrm{CI} = \hat{\theta} \pm z_{\alpha/2}\sigma_{\hat{\theta}}$, $\quad z_{\alpha/2}$ critical value
Small-sample CIs:
for μ: $\mathrm{CI} = \overline{Y} \pm t_{\alpha/2}(S/\sqrt{n})$, $\nu = df = n - 1$
for $\mu_1 - \mu_2$: $\mathrm{CI} = \overline{Y}_1 - \overline{Y}_2 \pm t_{\alpha/2}S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, $\nu = n_1 + n_2 - 2$ and **pooled sample estimator/variance** $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$;
S_p^2 **unbiased and consistent** est. of σ^2, $\mathbb{E}[S_p^2] = \sigma^2$, $\mathbb{V}\left[S_p^2\right] = \frac{2\sigma^4}{n_1+n_2-2} = \frac{\sigma^4}{n-1}$
CI for σ^2**:** $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}\right)$, $\nu = n$
Choosing sample size: Solve $z_{\alpha/2}\sigma_{\hat{\theta}} = B$ for n, with $B =$ desired bound
Show U **pivotal qty:** Show $F_U(u)$ indep of θ, e.g. $U = Y/\theta$, $F_U(u) = P(U \leq u) = P(Y \leq \theta u) = 2u - u^2 \quad 0 < u < 1$, indep of θ
Cdf to pdf: $F_Y(y) = \int_0^y f_Y(t)\mathrm{d}t$
Order statistics: $f_{(n)}(y) = nF(y)^{n-1}f(y)$
Efficiency: $\mathrm{eff}(\hat{\theta}_1, \hat{\theta}_2) = \mathbb{V}\left[\hat{\theta}_2\right]/\mathbb{V}\left[\hat{\theta}_1\right]$
Consistency: $\hat{\theta}$ unbiased and $\lim_{n\to\infty}\mathbb{V}\left[\hat{\theta}\right] = 0$. **Law of Large Numbers (LLN):** sample vals converges to pop. vals e.g. $\overline{Y} \to \mathbb{E}[Y]$. From LLN, sample vals are **consistent** ests. of pop. vals. $\hat{\theta}$ (op) $\hat{\theta}'$ **consistent** est. of θ (op) θ'. $\hat{\theta} \to \theta$ and $\hat{\theta}' \to \theta' \Rightarrow \hat{\theta} \pm \hat{\theta}' \to \theta \pm \theta', \hat{\theta} \times \hat{\theta}' \to \theta \times \theta', \hat{\theta}/\hat{\theta}' \to \theta/\theta'$ if $\theta' \neq 0$, $f(\hat{\theta}) \to f(\theta)$ if f real-valued fn continuous at θ.
Central Limit Theorem: $U_n = \frac{\sum_i^n Y - n\mathbb{E}[Y]}{\sqrt{\mathbb{V}[Y]n}} = \frac{\sum_i^n Y - n\mu}{\sigma\sqrt{n}} = \frac{\overline{Y} - \mathbb{E}[Y]}{\sqrt{\mathbb{V}[Y]/n}} \to \frac{(\overline{Y}-\mu)\sqrt{n}}{\sqrt{\sigma^2}}$ std Normal distr.. $W_n \to 1 \Rightarrow U_n/W_n \to$ std Normal distr.
Sufficiency: U sufficient for θ if $P(Y_1 = y_1, ..., Y_n = y_n|U = u) = P(Y_1 = y_1, ..., Y_n = y_n, U = u)/P(U = u)$ indep of θ
Factorization Criterion proving sufficiency: $L(\theta) = f$ or $p(y_1, ..., y_n|\theta) = \prod_i^n f(y_i|\theta)$ if iid $= g(u, \theta) \times h(y_1, ..., y_n)$, g fn of only u and θ, h not fn of θ
Rao-Blackwell Theorem: $\hat{\theta}$ unbiased est. of θ and $\mathbb{V}\left[\hat{\theta}\right] < \infty$ and U sufficient stat. for $\theta \Rightarrow \hat{\theta}^* = \mathbb{E}[\hat{\theta}|U]$, $\mathbb{E}[\hat{\theta}^*] = \theta, \mathbb{V}\left[\hat{\theta}^*\right] < \mathbb{V}\left[\hat{\theta}\right]$
Typically $\hat{\theta}^*$ is **MVUE** of θ
MVUE (minimum variance unbiased estimation): Some fn of sufficient U, $h(U)$, $\mathbb{E}[h(U)] = \theta \Rightarrow h(U)$ MVUE of θ
MLE: Solve $\frac{\partial L(\theta)}{\partial\theta} = 0$ for θ. $\overline{Y}, \sum(Y_i - \overline{Y})^2/n$ are MLEs of μ, σ^2
Elements of a Statistical Test: Type I err: accept H_a when H_a false or reject H_0 when $H_0 true$; type II err: reject H_a when true or accept H_0 when false. $P(\text{type I err}) = P(\text{reject } H_0|H_0 \text{ true}) = \alpha, P(\text{type II err}) = \beta = 1 - \alpha$.
Common Large Sample Tests: Test μ: $Z = \frac{\hat{\mu}-\mu}{\sigma_{\hat{\mu}}} = \frac{\overline{Y}-\mu}{S/\sqrt{n}} > z_\alpha$ or $z_{\alpha/2}$ if 2-tailed. Test diff of μs: $Z = \frac{\overline{Y}_1 - \overline{Y}_2 - (\mu_1-\mu_2)_{H_0}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} > z_\alpha$ or $z_{\alpha/2}$ if 2-tailed.
CI $= \hat{\mu} \pm z_\alpha(S/\sqrt{n})$.
p-Values (Attained Significance Levels): Minimum threshold for RR. z test statistic, p-value $= P(Z \leq z)$, reject H_0 if $\alpha >$ p-value. Do the opposite sign for opposite RR.
Testing Statistics Concerning Variances: $H_0: \sigma_1^2 = \sigma_2^2$, $H_a: \sigma_1^2 >, <, \neq \sigma_2^2$, $F = S_1^2/S_2^2 \sim F(n_1 - 1, n_2 - 1)$, test $>, < F_\alpha$ or $F_{\alpha/2}$ if 2-tailed.
Most powerful test: $H_0: \theta = \theta_0, H_a: \theta = \theta_a$. RR is given by $\frac{L(\theta_0)}{L(\theta_a)} < k$ from **Neyman-Pearson Lemma**.
Likelihood Ratio test: Likelihood ratio $\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})}$, Ω_0 params under H_0, Ω params in unrestricted space, $\hat{\Omega}_0, \hat{\Omega}$ MLEs of $L(\Omega_0), L(\Omega)$. $-2\ln\lambda \sim \chi^2(r_0 - r)$, r_0 # free vars that are fixed in Ω_0 e.g. $H_0: p_1 = p_2 = p_3 = p_4 \Rightarrow r_0 = 3$, r # free vars fixed in unrestricted space i.e. often 0. Reject H_0 if $-2\ln\lambda > \chi^2_\alpha$ crit. val., generally reject if $\lambda < k$
Linear models (least squares): Line $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$, $\hat{\beta}_1 = \frac{\sum(x_i-\overline{x})(y_i-\overline{y})}{\sum(x_i-\overline{x})^2} = \frac{S_{xy}}{S_{xx}} = \frac{\sum(x_i-\overline{x})y_i - \overline{y}\sum(x_i-\overline{x})}{S_{xx}} = \frac{\sum(x_i-\overline{x})y_i}{S_{xx}}$, $\hat{\beta}_0 = \overline{y} - \hat{\beta}_1\overline{x}$. $Y = \beta_0 + \beta_1 + \epsilon$, $\mathbb{V}[\epsilon] = \sigma^2$, $\mathbb{E}[\epsilon] = 0$, $\mathbb{V}[Y] = \mathbb{V}[\epsilon] = \sigma^2$. **More properties: 1)** $\hat{\beta}_0, \hat{\beta}_1$ unbiased for $\mathbb{E}[\beta_0], \mathbb{E}[\beta_1]$ **2)** $\mathbb{V}\left[\hat{\beta}_0\right] = c_{00}\sigma^2, c_{00} = \sum x_i^2/(nS_{xx})$ **3)** $\mathbb{V}\left[\hat{\beta}_1\right] = c_{11}\sigma^2 \Rightarrow S_{xx} = S^2/\mathbb{V}\left[\hat{\beta}_1\right], c_{11} = 1/S_{xx}$ **4)** $\mathrm{Cov}[\hat{\beta}_0, \hat{\beta}_1] = c_{01}\sigma^2, c_{01} = -\overline{x}/S_{xx}$ **5)** $\mathrm{SSE}/(n-2)$ unbiased for σ^2, $\mathrm{SSE} = S_{yy} - \hat{\beta}_1 S_{xy}, S_{yy} = \sum(y_i - \overline{y})^2$.
If $\epsilon_i \sim N$**: 6)** $\hat{\beta}_0, \hat{\beta}_1 \sim N$ **7)** $\frac{(n-2)S^2}{\sigma^2} \sim \chi^2(n-2)$ **8)** S^2 indep of $\hat{\beta}_0$ and $\hat{\beta}_1$.
Hypotheses for β_i**:** $H_0: \beta_i = \beta_{i0}$; $H_a: \beta_i >, <, \neq \beta_{i0}$; **test statistic** $T = \frac{\hat{\beta}_i - \beta_{i0}}{S\sqrt{c_i i}}$ or $\frac{\hat{\beta}_i - \beta_{i0}}{\sqrt{\mathbb{V}[\hat{\beta}_i]}}$, **RR:** $t > t_\alpha, < -t_\alpha, |t| > t_{\alpha/2}$ t_α based on $n-2$ df, $100(1-\alpha)\%$ **CI** for $\beta_i = \hat{\beta}_i \pm t_{\alpha/2}S\sqrt{c_{ii}}$. **DEF:** $Z \sim N(0, 1), W \sim \chi^2(\nu)$, both indep $\Rightarrow T = \frac{Z}{\sqrt{W/\nu}} \sim t(\nu)$.

Hypotheses for $\theta = a_o\beta_0 + a_1\beta_1$**:** $H_0 : \theta = \theta_0$; $H_a : \theta >, <, \neq \theta_0$; **test stat** $T = \frac{\hat{\theta}-\theta_0}{S\sqrt{\mathbb{V}[\hat{\theta}]/\sigma^2}}$, $\mathbb{V}\left[\hat{\theta}\right] = \left(\frac{a_0^2\frac{\sum x_i^2}{n}+a_1^2-2a_0a_1\overline{x}}{S_{xx}}\right)\sigma^2$; **RR:** $t > t_\alpha, < -t_\alpha, |t| > t_{\alpha/2}$, based on $n-2$ df, **CI** for $\theta : \hat{\theta} \pm t_{\alpha/2}S\sqrt{\mathbb{V}\left[\hat{\theta}\right]/\sigma^2}$, and for (mean of Y when $x = x^*$) $\mathbb{E}[Y] = \beta_0 + \beta_1 x^* : \hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2}S\sqrt{\frac{1}{n} + \frac{(x^*-\overline{x})^2}{S_{xx}}}$. **PI** for Y when $x = x^*$ is: $\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2}S\sqrt{1 + \frac{1}{n} + \frac{(x^*-\overline{x})^2}{S_{xx}}}$.
Correlation: Correl coef $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{xy}}} = \hat{\beta}_1\sqrt{\frac{S_{xx}}{S_{yy}}}$. **Coef of determination** $r^2 = \frac{\hat{\beta}_1 S_{xx}}{S_{yy}} = \frac{S_{yy}-SSE}{S_{yy}} = 1 - SSE/S_{yy} = \frac{SST}{SST+SSE}$
ANOVA: Table for one-way layout:

	df	SS	MS	F
Treatment	$k-1$	$SST = SS_{Reg} = \sum_i^k n_i(\overline{Y}_i - \overline{Y})^2 = \sum_i^k \overline{Y}_i^2/n_i - CM$	$MST = MS_{Reg} = \frac{SST}{k-1}$	$\frac{MST}{MSE}$
Error	$n-k$	$SSE = SS_{Res} = \sum_i^k\sum_j^{n_i}(Y_{ij} - \overline{Y}_i)^2 = \sum_i^k (n_i-1)S_i^2$	$MSE = MS_{Res} = \frac{SSE}{n-k}$	
Total	$n-1$	$TSS = \sum_i^k\sum_j^{n_i}(Y_i - \overline{Y})^2 = \sum_i^k\sum_j^{n_i} Y_{ij}^2 - CM$		

For lin reg, add 1 to k: $k-1 \to k, n-k \to n-(k+1)$.
$TSS = SST + SSE$; correction for the mean $CM = \frac{(\text{sum of all obs})^2}{n} = \frac{1}{n}\left(\sum_i^k\sum_j^{n_i} Y_{ij}\right)^2 = n\overline{Y}^2$; $S_i^2 = \frac{1}{n_i-1}\sum_j^{n_i}(Y_{ij} - \overline{Y})$; $S^2 = MSE = \frac{SSE}{(n_1-1)+...+(n_k-1)} = \frac{SSE}{n-k}$ (same as **pooled sample var**).
F test: $W_1 \sim \chi^2(\nu_1), W_2 \sim \chi^2(\nu_2), W_1 \perp W_2 \Rightarrow F = \frac{W_1/\nu_1}{W_2/\nu_2} \sim F(\nu_1, \nu_2)$. Test $H_0 : \mu_1 = ... = \mu_k$, **RR:** $F = \frac{MST}{MSE} > F_\alpha$.
Analysis of categorical data: χ^2 **test:** $X^2 = \sum_i^k \frac{(n_i - \mathbb{E}[n_i])^2}{\mathbb{E}[n_i]} = \sum_i^k \frac{(n_i - np_i)^2}{np_i} > \chi^2_\alpha(df = k-1)$, k # categories e.g. types of peas.
Contingency Table: n_{ij} observed freq row i col j, p_{ij} prob of obs falling into cell ij; MLEs $\hat{p}_{ij} = n_{ij}/n, \hat{p}_i = r_i/n$ (r_i # obs in row i), $\hat{p}_j = c_j/n, \hat{\mathbb{E}}[n_{ij}] = n(\hat{p}_i\hat{p}_j) = (r_ic_j)/n$; **test** $X^2 = \sum_{ij}\frac{n_{ij} - \hat{\mathbb{E}}[n_{ij}]}{\hat{\mathbb{E}}[n_{ij}]} > \chi^2_\alpha(df = (r-1)(c-1))$ with r, c nums rows and cols, then reject independence of the 2 classifications.

Distributions

Bernoulli(p)**:** $p(y) = p^y(1-p)^{1-y}$, $\mathbb{E}[Y] = p$, $\mathbb{V}[Y] = p(1-p)$ $\quad y = 0, 1$
Binomial(n, p)**:** $(a+b)^n = \sum_{j=0}^n \binom{n}{j}a^jb^{n-j}$, $p(y) = \binom{n}{y}p^y(1-p)^{n-y}$, $\mathbb{E}[Y] = np$, $\mathbb{V}[Y] = np(1-p)$ $\quad y = 0 : n$
Multinomial(n)**:** $p(y) = \frac{n!}{x_1!...x_n!}p_1^{x_1}...p_n^{x_n}$
Poisson(λ)**:** $p(y) = \frac{\lambda^y}{y!}e^{-\lambda}$, $\mathbb{E}[Y] = \mathbb{V}[Y] = \lambda$ $\quad y = 0, 1, ...;$ $\quad m(t) = \exp\left\{\lambda(e^t - 1)\right\}$
$\sum_i^n Y_i \sim$ Poisson($n\lambda$)
Power family(α, θ)**:** $f(y) = \alpha y^{\alpha-1}/\theta^\alpha$ $\quad 0 \leq y \leq \theta$, $\quad$ 0 otherwise, $F(y) = \frac{y^\alpha}{\theta^\alpha}$, $\mathbb{E}[Y] = \alpha\theta/(\alpha+1)$
Uniform(θ_1, θ_2)**:** $f(y) = \frac{1}{\theta_2 - \theta_1}$ $\quad y \in (\theta_1, \theta_2)$, $\quad$ 0 otherwise, $F(y) = \frac{y-\theta_1}{\theta_2-\theta_1}$, $\mathbb{E}[Y] = \frac{\theta_1+\theta_2}{2}$, $\mathbb{V}[Y] = \frac{(\theta_2+\theta_1)^2}{12}$
Gamma(α, β)**:** $f(y) = \frac{1}{\beta^\alpha\Gamma(\alpha)}y^{\alpha-1}e^{-y/\beta}$ if $y \geq 0$, 0 otherwise. $X \sim \Gamma(\alpha_1, \beta)$ and $Y \sim \Gamma(\alpha_2, \beta) \Rightarrow X + Y \sim \Gamma(\alpha_1 + \alpha_2, \beta)$. $X \sim \Gamma(\alpha, \beta) \Rightarrow 2X/\beta \sim \chi^2(2\alpha)$
$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}\mathrm{d}t$, $\quad \Gamma(n) = (n-1)\Gamma(n-1) = (n-1)!$, $\quad \Gamma(1/2) = \sqrt{\pi}$
$\mathbb{E}[Y] = \alpha\beta$, $\quad \mathbb{V}[Y] = \alpha\beta^2$, $\quad \mathbb{E}[Y^2] = \alpha(\alpha+1)\beta^2$, $\mathbb{E}[Y^3] = \alpha(\alpha+1)(\alpha+2)\beta^2$, ... $\quad m(t) = \left(\frac{1}{1-\beta t}\right)^\alpha$

Chi-squared(ν)**:** $\alpha = \nu/2$, $\beta = 2$, $\nu = 1, 2, ...$ $\mathbb{E}[Y] = \nu, \mathbb{V}[Y] = 2\nu$. $X \sim \chi^2(a), Y \sim \chi^2(b) \Rightarrow X + Y \sim \chi^2(a+b), X - Y \sim \chi^2(a-b)$
Exponential(β)**:** $\alpha = 1$, standard $\beta = 1$, $f(y) = \frac{1}{\beta}e^{y/\beta}$, $F(y) = 1 - e^{-y/\beta}$, $\mathbb{E}[Y] = \beta$, $\mathbb{V}[Y] = \beta^2$
Normal(μ, σ^2)**:** $f(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}$, $F(y) = \int_{-\infty}^y f(t)\mathrm{d}t$, $\mathbb{E}[Y] = \mu$, $\mathbb{V}[Y] = \sigma^2$. $Y_i \sim N(\mu, \sigma^2) \Rightarrow \frac{\overline{Y}-\mu}{\sigma} \sim N(0, 1)$ (std normal), $(n-1)S^2/\sigma^2 = \sum(Y_i - \overline{Y})^2/\sigma^2 \sim \chi^2(n-1)$
Standard Normal($\mu = 0, \sigma^2 = 1$)**:** $f(y) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2}\right\}$. $Y \sim N(0, 1) \Rightarrow Y^2 \sim \chi^2(1)$
Beta: $f(y) = \frac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1}$, $y \in [0, 1]$, 0 otherwise; $\quad B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\mathrm{d}y$
$\alpha = \beta = 1 \Rightarrow Y \sim$ Uniform(0, 1)
Weibull(α, β)**:** $f(y) = \alpha\beta y^{\alpha-1}e^{-\beta y^\alpha}$ $\quad y \geq 0$, $\quad$ 0 otherwise, $F(y) = 1 - e^{-\beta y^\alpha}$, $\mathbb{E}[Y] = \frac{\Gamma(1+1/\alpha)}{\beta^{1/\alpha}}$, $\mathbb{V}[Y] = \frac{\Gamma(1+2/\alpha) - \Gamma(1+1/\alpha)^2}{\beta^{2/\alpha}}$

Examples

1.a. $Y_i \sim N(\mu, \sigma = 1)$. **Show MVUE of** μ^2 **is** $\hat{\mu}^2 = \overline{Y}^2 - 1/n$**:** $L(\mu, \sigma^2) = (2\pi)^{-n/2}\exp\left\{-\sum(Y_i - \mu)^2/2\right\} = (2\pi)^{-n/2}\exp\left\{-\sum(Y_i - \overline{Y} + \overline{Y} - \mu)^2/2\right\} = (2\pi)^{-n/2}\exp\left\{-\frac{1}{2}\left[\sum(Y_i - \overline{Y})^2 - 2(\overline{Y} - \mu)\sum(Y_i - \overline{Y}) + \sum(\overline{Y} - \mu)^2\right]\right\} = (2\pi)^{-n/2}\exp\left\{-\frac{1}{2}\left[\sum(Y_i - \overline{Y})^2 + n(\overline{Y} - \mu)^2\right]\right\}$
$= \left[\exp\left\{n(\overline{Y} - \mu)^2\right\}\right]\left[(2\pi)^{-n/2}\exp\left\{\sum(Y_i - \overline{Y})^2/2\right\}\right] = g(\overline{Y}, \mu)h(Y_i)$. From **fact. crit.**, $\overline{Y}$ **sufficient** for μ. Want fn of $\overline{Y}$ st $\mathbb{E}[f(\overline{Y})] = \mu^2$ (i.e. unbiased).
$\mathbb{E}[\overline{Y}^2] = \mu + \sigma^2/n = \mu^2 + 1/n \Rightarrow \mathbb{E}[\overline{Y}^2 - 1/n] = \mu^2 \Rightarrow \overline{Y}^2 - 1/n$ **MVUE** of μ^2. ■
1.b. Find variance of $\hat{\mu}^2$**:** $\mathbb{V}\left[\overline{Y}^2 - 1/n\right] = \mathbb{V}\left[\overline{Y}^2\right] = \mathbb{V}\left[\frac{1}{n}\left[\sqrt{n}(\overline{Y} - \mu + \mu)\right]^2\right] = \frac{1}{n^2}\mathbb{V}\left[(\sqrt{n}(\overline{Y} - \mu) + \mu\sqrt{n})^2\right] = \frac{1}{n^2}\mathbb{V}\left[Z^2 + 2Z\mu\sqrt{n} + n\mu^2\right] = \frac{1}{n^2}\left(\mathbb{V}\left[Z^2\right] + 4n\mu^2\mathbb{V}[Z] + 2\mu\sqrt{n}\mathrm{Cov}[Z^2, Z]\right)$. $Z^2 \sim \chi^2(1) \Rightarrow \mathbb{V}\left[Z^2\right] = 2\nu = 2$, $\mathbb{V}[Z] = 1, \mathrm{Cov}[Z^2, Z] = 0$. Therefore, $\mathbb{V}\left[\overline{Y}^2 - 1/n\right] = \frac{2+4n\mu^2}{n^2}$.
2. Rewrite test stat: $T = \frac{\hat{\beta}_1 - 0}{S/\sqrt{S_{xx}}} = \frac{\hat{\beta}_1\sqrt{n-2}}{\sqrt{S_{yy} - \hat{\beta}_1 S_{xy}/\sqrt{S_{xx}}}} = \frac{\hat{\beta}\sqrt{n-2}\sqrt{S_{xx}}}{\sqrt{S_{yy} - \hat{\beta}_1 S_{xy}}} = \frac{r\sqrt{n-2}}{\sqrt{1 - \hat{\beta}_1\frac{S_{xy}}{S_{yy}}}} = \frac{r\sqrt{n-2}}{\sqrt{1 - \frac{S_{xy}}{S_{xx}}\frac{S_{xy}}{S_{yy}}}} = \frac{r\sqrt{n-2}}{1 - \frac{\sqrt{S_{xy}^2}}{S_{xx}S_{yy}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$.
3.a. Prove $SSE/\sigma^2 \sim \chi^2((n_1-1) + ... + (n_k-1) = n-k)$**:** $\frac{(n_i-1)S_i^2}{\sigma^2} \sim \chi^2(n_i - 1) \Rightarrow \sum_i^k \frac{SSE}{\sigma^2} \sim \chi^2(n-k)$. ■
3.b. Prove $TSS/\sigma^2 \sim \chi^2(n-1)$ **under** $H_0 : \mu_1 = ... = \mu_k$**:** $TSS = \sum_i^k\sum_j^{n_i}(Y_{ij} - \overline{Y})^2 = \sum_{ij}(Y_{ij} - \overline{Y})^2$, $\frac{TSS}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. ■
3.c. Prove $SST/\sigma^2 \sim \chi^2(k-1)$ **under** H_0**:** $\frac{TSS}{\sigma^2} \sim \chi^2(n-1)$, $\frac{SSE}{\sigma^2} \sim \chi^2(n-k), SST = TSS - SSE \Rightarrow \frac{SST}{\sigma^2} = \frac{TSS}{\sigma^2} - \frac{SSE}{\sigma^2} \sim \chi^2(n-1-(n-k) = k-1)$.
■
3.d. Prove $F = MST/MSE \sim F(k-1, n-k)$ **under** H_0**:** $SST/\sigma^2 \sim \chi^2(k-1), SSE/\sigma^2 \sim \chi^2(n-k) \Rightarrow F = \frac{(SST/\sigma^2)/(k-1)}{(SSE/\sigma^2)/(n-k)} = \frac{SST/(k-1)}{SSE/(n-k)} = \frac{MST}{MSE} \sim F(k-1, n-k)$. ■
4. Get S^2 **from std err of mean:** std err $= S/\sqrt{n} \Rightarrow S^2 = n(\text{std err})^2$