Math 341 Formula Sheet

$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2)$. $F(y_1, y_2) = P(Y_1 \le y_1, Y_2 \le y_2)$.

$F(y_1, y_2) = \int_{-\infty}^{y_1}\int_{-\infty}^{y_2} f(t_1, t_2)\,dt_2\,dt_1$. If $y_1^* \ge y_1$ and $y_2^* \ge y_2$, then

$F(y_1^*, y_2^*) - F(y_1^*, y_2) - F(y_1, y_2^*) + F(y_1, y_2) \ge 0$.

Note that $f(y_1, y_2) \ge 0$ and $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y_1, y_2)\,dy_1 dy_2 = 1$.

$p_1(y_1) = \sum_{\text{all } y_2} p(y_1, y_2)$ and $p_2(y_2) = \sum_{\text{all } y_1} p(y_1, y_2)$.

$f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)\,dy_2$ and $f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2)\,dy_1$.

$p(y_1 | y_2) = P(Y_1 = y_1 | Y_2 = y_2) = \dfrac{p(y_1, y_2)}{p_2(y_2)}$. $F(y_1 | y_2) = P\{Y_1 \le y_1 | Y_2 = y_2\}$.

$f(y_1 | y_2) = \dfrac{f(y_1, y_2)}{f_2(y_2)}$, $f(y_2 | y_1) = \dfrac{f(y_1, y_2)}{f_1(y_1)}$. $F(y_1, y_2) = F_1(y_1)F_2(y_2)$.

$p(y_1, y_2) = p_1(y_1)p_2(y_2)$. $f(y_1, y_2) = f_1(y_1) f_2(y_2)$.

$f(y_1, y_2) = g(y_1)h(y_2), g \ge 0, h \ge 0$.

$E[g(Y_1, Y_2, ..., Y_k)] = \sum_{\text{all } y_k} \cdots \sum_{\text{all } y_1} g(y_1, y_2, ..., y_n)\, p(y_1, y_2, ..., y_k)$.

$E[g(Y_1, Y_2, ..., Y_k)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(y_1, y_2, ..., y_k)\, f(y_1, y_2, ..., y_k)\, dy_1, dy_2, ..., dy_k$.

$cov(Y_1, Y_2) = E\left[\left(Y_1 - E(Y_1)\right)\left(Y_2 - E(Y_2)\right)\right]$..

$\text{Var}\left(aY_1 + bY_2\right) = a^2\text{Var}(Y_1) + b^2\text{Var}(Y_2) + 2ab\text{Cov}(Y_1, Y_2)$. $\text{Cov}(Y_1, Y_2) = E[Y_1Y_2] - \mu_1\mu_2$.

$Cov(\sum_{i=1}^{n} a_iY_i, \sum_{j=1}^{m} b_jX_j,) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_ib_jCov(Y_i, X_j)$

$\text{Var}\left(\sum_{i=1}^{n} a_iY_i\right) = \sum_{i=1}^{n} a_i^2\text{Var}(Y_i) + 2\sum_{1\le i<j\le n}\sum a_ia_j\text{Cov}(Y_i, Y_j)$. $\rho = \dfrac{\text{Cov}(Y_1, Y_2)}{\sigma_1\sigma_2}$.

If $h(y)$ is either increasing or decreasing for all y such that $f_Y(y) > 0$, then $U = h(Y)$ has density function $f_U(u) = f_Y\left[h^{-1}(u)\right]\left|\dfrac{dh^{-1}(u)}{du}\right|$.

Suppose that Y_1 and Y_2 are continuous random variables with joint density function $f_{Y_1,Y_2}(y_1, y_2)$ and that for all (y_1, y_2) such that $f_{Y_1,Y_2}(y_1, y_2) > 0$,
$U_1 = h_1(Y_1, Y_2)$ *and* $U_2 = h_2(Y_1, Y_2)$ is a one-to-one transformation from (y_1, y_2) to (u_1, u_2)

with inverse $y_1 = h_1^{-1}(u_1, u_2)$, $y_2 = h_2^{-1}(u_1, u_2)$. If $h_1^{-1}(u_1, u_2)$ and $h_2^{-1}(u_1, u_2)$ have continuous partial derivatives with respect to u_1 and u_2 and Jacobian

$$J = \det\begin{bmatrix} \frac{\partial h_1^{-1}}{\partial u_1} & \frac{\partial h_1^{-1}}{\partial u_2} \\ \frac{\partial h_2^{-1}}{\partial u_1} & \frac{\partial h_2^{-1}}{\partial u_2} \end{bmatrix}$$, then the joint density

$f_{U_1,U_2}(u_1, u_2) = f_{Y_1,Y_2}\left(h_1^{-1}(u_1, u_2), h_2^{-1}(u_1, u_2)\right)|J|$, where $|J|$ is the absolute value of the Jacobian.

Let $Y_{(k)}$ represent the k^{th} order statistics then its density function can be described by $g_{(k)}(y_k) = \frac{n!}{(k-1)!(n-k)!}[F(y_k)]^{k-1}[1-F(y_k)]^{n-k} f(y_k)$. Let $Y_{(j)}$ represent the j^{th} order statistics and $Y_{(k)}$ represent the k^{th} order statistics ($j < k$) then their joint density function can be described by

$$g_{(j)(k)}(y_j, y_k) = \frac{n!}{(j-1)!(k-j-1)!(n-k)!}[F(y_j)]^{j-1}[F(y_k)-F(y_j)]^{k-j-1}[1-F(y_k)]^{n-k} f(y_j) f(y_k), y_j < y_k.$$

$$S^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}, \quad {}_N\mathrm{P}_n = \frac{N!}{(N-n)!}, \quad \binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

Let $Y_1, \ldots, Y_n$ be independent identically distributed with mean $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$.

Define $U_n = \frac{\sum_{i=1}^{n} Y_i - n\mu}{\sigma\sqrt{n}} = \frac{\frac{\sum_{i=1}^{n} Y_i}{n} - \mu}{\left(\sigma / \sqrt{n}\right)}$. Then the distribution function of U_n converges to the standard normal distribution function as $n \to \infty$. Let $Y_1, \ldots, Y_n$ be a random sample from N(μ, σ^2) then $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{\sigma^2}$ has distribution of χ^2, d.f. = $n - 1$ and $\bar{Y}$ and S^2 are independent.

Binomial $p(y) = \binom{n}{y} p^y (1-p)^{n-y}$, y = 0, 1, …, n, V(Y) = np(1-p)

Poisson $P\{Y = i\} = e^{-\lambda}\frac{\lambda^i}{i!}, \; i = 0, 1, 2, \ldots., \quad E[Y] = Var(Y) = \lambda$

Negative Binomial $P\{Y = n\} = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \; n = r, r+1, \ldots., \quad E[Y] = \frac{r}{p},$

$Var(Y) = \frac{r(1-p)}{p^2}$. $m(t) = \left[\frac{pe^t}{1-(1-p)e^t}\right]^r$, $t < -\ln(1-p)$.

Normal density $f(y) = \frac{1}{(\sqrt{2\pi})\sigma} e^{-(y-\mu)^2/2\sigma^2}$, $m(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$,

Gamma density $f(y) = \begin{cases} \frac{y^{\alpha-1} e^{-\frac{y}{\beta}}}{\Gamma(\alpha)\beta^\alpha}, & if \quad y \geq 0 \\ 0, & if \quad y < 0 \end{cases}$, $E[Y] = \alpha\beta, Var(Y) = \alpha\beta^2$.

$m(t) = \frac{1}{(1-\beta t)^\alpha}$, $t < \frac{1}{\beta}$. $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$, $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$, $\Gamma(n) = (n-1)!$,

$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Beta density

$$f(y) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}, 0 < y < 1 \\ 0, \qquad elsewhere. \end{cases} \quad E[Y] = \frac{a}{a+b}, Var(Y) = \frac{ab}{(a+b+1)(a+b)^2}.$$

Uniform (a, b), $Var(Y)$ = $(b-a)^2/12$.

$m_U(t) = E\left(e^{tU}\right)$, $E[U^n] = \frac{d^n}{dt^n} m_U(t)\Big|_{t=0}$, $m(t_1, t_2) = E\, e^{t_1 Y_1 + t_2 Y_2}$,

$E(Y_1^i Y_2^j) = \frac{\partial^{i+j}}{\partial t_1^i \partial t_2^j} m(t_1, t_2)\Big|_{t_1 = t_2 = 0}$.

Tchebysheff's theorem: Let k a known constant >1 P (|Y- E(Y)| ≤ k σ_Y) ≥ 1 - $\frac{1}{k^2}$.

Bivariate Normal density $f(y_1, y_2) = \frac{e^{-Q/2}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$,

$$Q = \frac{1}{1-\rho^2}\left[\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{y_1-\mu_1}{\sigma_1}\right)\left(\frac{y_2-\mu_2}{\sigma_2}\right) + \left(\frac{y_2-\mu_2}{\sigma_2}\right)^2\right].$$

MSE($\hat{\theta}$) = Var ($\hat{\theta}$) + (Bias[$\hat{\theta}$])^2.

Target Parameter θ	Sample Size(s)	Point Estimator $\hat{\theta}$	E($\hat{\theta}$)	Standard Error $\sigma_{\hat{\theta}}$
μ	n	$\bar{Y}$	μ	$\frac{\sigma}{\sqrt{n}}$
p	n	$\hat{p} = \frac{Y}{n}$	p	$\sqrt{\frac{pq}{n}}$

$\mu_1 - \mu_2$	n_1 *and* n_2	$\bar{Y}_1 - \bar{Y}_2$	$\mu_1 - \mu_2$	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$p_1 - p_2$	n_1 *and* n_2	$\hat{p}_1 - \hat{p}_2$	$p_1 - p_2$	$\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

$Z = \frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}}$. If the parameter θ is $\mu, p, \mu_1 - \mu_2,$ or $p_1 - p_2$ then the two sided (1 - α) CI may be given by $\hat{\theta} \pm z_{\alpha/2}\sigma_{\hat{\theta}}$. The small sample $1-\alpha$ confidence interval for $\mu_1 - \mu_2$, with equal but unknown variances and additional assumptions is given by

$\bar{Y}_1 - \bar{Y}_2 \pm t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, where $S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$. $n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_a - \mu_0)^2}$. $T = \frac{\bar{Y} - \mu}{S/\sqrt{n}}$,

$\bar{Y} \pm t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)$. ***Factorization criterion***: Let U be a statistic based on the random sample $Y_1, \ldots, Y_n$. Then U is a ***sufficient statistic*** for the estimation of a parameter θ if and only if the likelihood $L(\theta) = L(y_1, \ldots, y_n | \theta)$ can be factored into two nonnegative functions, $L(y_1, \ldots, y_n | \theta) = g(u, \theta) \times h(y_1, \ldots, y_n)$ where $g(u, \theta)$ is a function only of u and θ and $h(y_1, \ldots, y_n)$ is not a function of θ.

Total Sum of Square (SS) = $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y})^2$, Total SS = SST +SSE.

SST= $\sum_{i=1}^{k} n_i(\bar{Y}_{i\bullet} - \bar{Y})^2$ and SSE = $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{i\bullet})^2$.

ANOVA table for a one-way layout.

Source	df	SS	MS	*F*
Treatments	***k - 1***	**SST**	$\text{MST} = \frac{\text{SST}}{k-1}$	$\frac{\text{MST}}{\text{MSE}}$
Error	***n - k***	**SSE**	$\text{MSE} = \frac{\text{SSE}}{n-k}$	
Total	***n - 1***	$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y})^2$		

$\chi^2 = \sum_{i=1}^{k} \frac{[n_i - E(n_i)]^2}{E(n_i)} = \sum_{i=1}^{k} \frac{[n_i - np_i]^2}{np_i}$.