Conditional variance identity $Var[X] = Var(E[X|Y] + E(Var[X|Y]))$
Law of iterated Expectation, $E(E[X|Y]) = E[X] = \mu_X$.
Cauchy-Schwarz inequality $|EXY| \le E|XY| \le \sqrt{E|X|^2 E|Y|^2}$
$\text{Cov(X,Y)}=|\text{E}(\text{X}-\mu_X)(\text{Y}-\mu_Y)| \le \sqrt{\text{E}(\text{X}-\mu_X)^2\text{E}(\text{Y}-\mu_Y)^2} = \sqrt{\text{VarXVarY}}$

Proposition 1.3.1 $MSE(T(X)) = Bias^2(T(X)) + Var(T(X))$
Unbiased Estimator Bias(θ)=$E_\theta[T(X)] - q(\theta)$
$\{P_\eta : \eta \in \varepsilon\}$ be a k-dim **canonical exp. family(*)** with s.s. $T(X)$
$q(\mathbf{x}, \eta) = h(x)\exp[\eta T(x) - A(\eta)]$, $x \in \mathcal{X} \subset \mathbb{R}^q$ is free from para,
where $A(\eta) = \log(\int \cdots \int h(x)\exp[\eta T(x)]dx)$ (similar to m.g.f.)
and $\varepsilon = \{\eta : A(\eta) < \infty\}$ is the natural para space;
Facotrization Theorem $T(X)$ is sufficient.
$\varepsilon \neq \emptyset$, then $T(X)$ is complete.

Theorem 1.6.2, X is distributed in form (*), $\eta(\theta) \implies A(\eta)$;
η is an interior point of the natural parameter space ε.
the moment-generating function of T(X) exist and is given by
$M_X(t) = \exp[A(t+\eta) - A(\eta)]$; $E(T(\underline{X})) = A'(\eta)$; $Var(T(\underline{X})) = A''(\eta)$

Rao-Blackwell Theorem$X \sim \{P_\theta : \theta \in \Theta\}$, $T(X)$ is a sufficient statistic for θ, $S(x)$ is a unbiased estimator for $q(\theta)$. function of T.
Then $\tilde{S}(x) = E_\theta[S(x)|T(X)]$ is unbiased for$q(\theta)$ and
$Var_\theta \tilde{S}(x) \le Var_\theta S(x)$, $\forall \theta \in \Theta$ *not proof unique*

Completemess if the solution to $E_\theta[g(T(X))] = 0$, $\theta \in \Theta$ is $g(T(X)) = 0$ a.s., then T(X) is complete
UMVUE if $E_\theta[T] = \mu(\theta)$ and $Var_\theta[T]) = \inf\{Var_\theta[S|E_\theta[S] = \mu(\theta)\}$
Then T is uniformly minimum variance unbiased estimator.
Lehmann-Scheffe theorem if T is a complete sufficient statistic $E_\theta[h(T)]) = \mu(\theta)\}$. Then, $h(T)$ is the UMVUE of $\mu(\theta)$.

Regularity Assumptions on the family $\{P_\theta : \theta \in \Theta\}$:
(I) The set $A = \{x : p(x, \theta) > 0\}$ does not depend on θ, $\forall x \in A, \theta \in \Theta$
score function $\frac{\partial}{\partial\theta}\log p(X, \theta)$ exists and is finite.
(II) If T is any statistic such that $E\theta(|T|) < \infty, \theta \in \Theta$, then the operations of integration and differentiation by θ can be interchanged in
$\frac{\partial}{\partial\theta}E\theta(T(X))) = \frac{\partial}{\partial\theta}\int T(x)p(x,\theta)dx = \int T(x)\frac{\partial}{\partial\theta}p(x,\theta)dx$

(Proposition 3.4.1.) $\{P\theta\}$ is an exponential family and $\eta(\theta)$ has a nonvanishing continuous derivative on Θ, then (I) and (II) hold.
$I(\theta) = E\left([\frac{\partial}{\partial\theta}\log p(x,\theta)]^2\right) = \int \left([\frac{\partial}{\partial\theta}\log p(x,\theta)]^2\right)p(x,\theta)dx$

Theorem 3.4.1.Cramer-Rao Inequality $\{P_\theta : \theta \in \Theta\}$ has density $p(x, \theta)$,$x \in A \subseteq \mathbb{R}^q$, $E_\theta(T(X)) = \psi(\theta)$ is differentiable$\forall\theta$
Suppose that I and II hold $0 < I(\theta) < \infty$ $\forall\theta$, $Var_\theta(T(X)) \ge \frac{[\psi'(\theta)]^2}{I(\theta)}$
Corollary 3.4.1 if T is an unbiased estimate of θ, $Var_\theta(T(X)) \ge \frac{1}{I(\theta)}$

Proposition 3.4.2 $X_1, \cdots, X_n$ is a sample from a popu with density,
$I(\theta) = nI_1(\theta)$; $Var_\theta(T(X)) \ge \frac{[\psi'(\theta)]^2}{nI_1(\theta)}$; $[\psi(\theta)]^T_{1\times k} I(\theta)^{-1}_{k\times k}[\psi'(\theta)]_{k\times 1}$

Theorem 3.4.2. $\{P_\theta : \theta \in \Theta\}$ satisfies assumptions (I) and (II).
There exists u.b. est. T^* of $\psi(\theta)$, which achieves CRLB $\forall\theta$.
Then $\{P\theta\}$ is a one-parameter exponential family of the form $p(x,\theta) = h(x)\exp\{\eta(\theta)T(x) - B(\theta)\}$
Conversely, if $\{P_\theta\}$ a one-para exp family of the form (*) with n.s.s. $T(X)$.
$\eta(\theta)$ has a continuous nonvanishing derivative on θ,
Then $T(X)$ achieves CRLB and is a UMVUE of $E_\theta(T(X))$

CRLB obtained from Q_η evaluated at $\eta = h(\theta)$ is the same as the bound obtained from P_θ
$\frac{\partial}{\partial\eta}\log q(x,\eta) = \left[\frac{\partial}{\partial\theta}\log p(x,\theta)\right]\frac{\partial\theta}{\partial\eta}, \frac{\partial}{\partial\eta}E_\eta(T(X)) = \left(\frac{\partial}{\partial\theta}\psi(\theta)\right)\frac{\partial\theta}{\partial\eta},$
where $\psi(\theta) = E_\theta(T(X))$. $Q_\eta = P_\theta \implies \frac{\partial}{\partial\eta}Q_\eta = \frac{\partial}{\partial\theta}P_\theta\frac{\partial\theta}{\partial\eta}$
$\frac{\left[\frac{\partial}{\partial\eta}E_\eta(T(X))\right]^2}{I(\eta)} = \frac{\left[\frac{\partial}{\partial\theta}E_\theta(T(X))\right]^2}{I(\theta)}$

Conv $X_n \xrightarrow{\mathcal{D}} X$ if $F_n(x) \to F(x)$ $\forall$ continuity points of F
Theorem1 Scheffe's Thm if X_n has a pdf $f_n(x) \to f(x), \forall x \in$ support. $f(x)$ is a pdf of a r.v x, then $f_n(x) \xrightarrow{\mathcal{D}} X$
Theorem2 $X_n \xrightarrow{\mathcal{D}} X$ iff $E[f_n(x)] \to E[f(x)]$, $f \in C$
C is the set of bounded and cont. fn.s

Conv $X_n \xrightarrow{\mathcal{P}} X$ iff $P(|X_n - c| > \varepsilon) \to 0$ $\varepsilon > 0$ as $n \to \infty$

Theorem $X_n \xrightarrow{\mathcal{P}} c$, iff $X_n \xrightarrow{\mathcal{D}} c$ degenerating.

Theorem if $h : \mathbb{R}^k \to \mathbb{R}^p$ is cont. in $c \subset \mathbb{R}^k$, $X_n \xrightarrow{\mathcal{P}} c$, Then $h(X_n) \xrightarrow{\mathcal{P}} h(c)$

Theorem $X_n \xrightarrow{\mathcal{P}} X$, $C_n \to c$ Then $C_n X_n \xrightarrow{\mathcal{P}} cX$

Theorem $X_n \xrightarrow{\mathcal{P}} X$, $Y_n \xrightarrow{\mathcal{P}} b$ Then $X_n + Y_n \xrightarrow{\mathcal{P}} X + b$

Cont. Mapping Theorem $X_n \xrightarrow{\checkmark} X$, g is a cont. fn. Then $g(X_n) \xrightarrow{\mathcal{P},\mathcal{D}} g(X)$

Slutsky's Theorem $X_n \xrightarrow{\mathcal{D}} X$, $Y_n \xrightarrow{\mathcal{P}} c$ $(X_n, Y_n) \xrightarrow{\mathcal{D}} (X, c)$
Corollary $X_n + Y_n \xrightarrow{\mathcal{D}} X + c$, $c \neq 0$ $\frac{X_n}{Y_n} \xrightarrow{\mathcal{D}} \frac{X}{c}$

Continuity Theorem

Weak Law of Large Numbers

Consistent a sequence of r.v.'s $\{X_n\}$, $X_n \xrightarrow{\mathcal{P}} \theta$, all $\theta \in \Theta$
Then $\{X_n\}$ is consitent for θ

Asym
Central Limit Theorem $X_1, ..X_n \overset{iid}{=} \mu, \sigma^2 < \infty$
$\frac{\sum X - n\mu}{\sqrt{n}\sigma} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xrightarrow{\mathcal{D}} N(0,1)$

Delta Method $\{X_n\}$ a sequence of estimators $\sqrt{n}(X_n - \mu) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$
Let f is a continuous differentiable fn at μ $f'(\mu) \neq 0$
$\sqrt{n}(f(X_n) - f(\mu)) \xrightarrow{\mathcal{D}} N(0, \sigma^2[f'(\mu)]^2)$

The frequency substitution estimates $h(\frac{\mathbf{N}}{n})$ of θ satisfies $h(\mathbf{p}(\theta)) = \theta$ for all $\theta \in \Theta$ where $\mathbf{p}(\theta) = (p(x_0, \theta), .., p(x_k, \theta))^T$
Many $h(\mathbf{p}(\theta))$ exist if $K > 1$. h is differentiable.
$\sqrt{n}\left(h(\frac{\mathbf{N}}{n}) - h(\mathbf{p}(\theta))\right) = \sqrt{n}\sum_{j=1}^3 \frac{\partial h}{\partial p_j}(\mathbf{p}(\theta))\left(\frac{N_j}{n} - p(x_j, \theta)\right) + \mathcal{O}_p(1)$
Theorem 5.4.1 $\sqrt{n}(h(\frac{\mathbf{N}}{n}) - \theta) \xrightarrow{\mathcal{D}} N(0, \sigma^2(\theta, h)$
Delta Method $\sqrt{n}\left(h(\frac{\mathbf{N}}{n}) - h(\mathbf{p}(\theta))\right)$ is asym normal with mean 0
asymptotic variance $\sigma^2(\theta, h) =$
$Var_\theta\left[\sum_{j=1}^k \frac{\partial h}{\partial p_j}(\mathbf{p}(\theta))\mathbf{1}_{(x_1 = x_j)}\right] = \sum_{j=1}^k \left[\frac{\partial h}{\partial p_j}(\mathbf{p}(\theta))\right]^2 p(x_j, \theta) - \left[\sum_{j=1}^k \frac{\partial h}{\partial p_j}(\mathbf{p}(\theta))p(x_j, \theta)\right]^2$

Multivariate CLT $X_1, ..X_n \overset{iid}{=} (\mu, \sigma^2)$, μ_4 is finite.
$\sqrt{n}((\overline{X_n}, \overline{X_n^2}) - (\mu_1, \mu_2)) \xrightarrow{\mathcal{D}} N(0, \Sigma)$
where $\Sigma = \begin{bmatrix} Var(X_1) & Cov(X_1, X_1^2) \\ Cov(X_1, X_1^2) & Var(X_1^2)\end{bmatrix} = \begin{bmatrix}\mu_2 - \mu_1^2 & \mu_3 - \mu_1\mu_2 \\ \mu_3 - \mu_1\mu_2 & \mu_4 - \mu_2^2\end{bmatrix}$
$J(\mu_1, \mu_2) = (\frac{\partial f}{\partial t_1}, \frac{\partial f}{\partial t_1})\Big|_{(t_1, t_2) = (\mu_1, \mu_2)} = [-2\mu_1 \quad 1]$, $f(t_1, t_2) = t_2 - t_1^2$
$J\Sigma J' = [-2\mu_1 \quad 1]\begin{bmatrix}\mu_2 - {\mu_1}^2 & \mu_3 - \mu_1\mu_2 \\ \mu_3 - \mu_1\mu_2 & \mu_4 - {\mu_2}^2\end{bmatrix}\begin{bmatrix}-2\mu_1 \\ 1\end{bmatrix} = \mu_4 - {\mu_2}^2 - 2(\mu_3 - \mu_1\mu_2) + \mu_2 - {\mu_1}^2$
Let $S_n^2 = \frac{n}{n-1}f(\overline{X_n}, \overline{X_n^2})$, where $f(x,y) = y - x^2$
By Delta Method,
$$\sqrt{n}\left(f(\overline{X_n}, \overline{X_n^2}) - f(\mu_1, \mu_2)\right) \xrightarrow{\mathcal{D}} N(0, J(\mu_1, \mu_2)\Sigma J')$$
$\sqrt{n}\left(S_n^2 - \sigma^2\right) \xrightarrow{\mathcal{D}} N(0, \eta^2)$,$\eta^2 = Var[-2\mu_1 X + X^2]$
Under the Cramer-Rao conditions for asymptotic normality, $\sqrt{n}(\hat{p}_{MLE} - p) \xrightarrow{\mathcal{L}} N(0, I(p)^{-1}) = N(0, p(1-p))$

Edgeworth Expansion

Lindberg CLT

$F_X(x|\theta_1, \theta_2) = \frac{x - \theta_1}{\theta_2 - \theta_1}$
$Y \sim U(0,1)$, $f_Y(y) = 1$, $F_Y(y) = y$, $Y_{(1)} \sim Beta(1, n)$, $Y_{(n)} \sim Beta(n, 1)$.
$f_{Y_{(1)}}(y) = \frac{n!}{(1-1)!(n-1)!}f_Y(y)[F_Y(y)]^{1-1}[1 - F_Y(y)]^{n-1} = n(1-y)^{n-1}$, $0 \le y \le 1$
$f_{Y_{(n)}}(y) = \frac{n!}{(n-1)!(n-n)!}f_Y(y)[F_Y(y)]^{n-1}[1 - F_Y(y)]^{n-n} = ny^{n-1}$, $0 \le y \le 1$