Chapter 3: Large Sample Theory (JAW Ch 2)

3.1 Modes of convergence (JAW 2.3, Ferg Ch 1)

(i) Defn: $X_n \to X$ $a.s.$ **if** $X_n(\omega) \to X(\omega)$ $\forall \omega \in A$ **where** $P(A^c) = 0$.

(ii) Equivalently (by defn of cgce of functions):
$\forall \epsilon > 0$, $P(\sup_{m \geq n} |X_m - X| > \epsilon) \to 0$ **as** $n \to \infty$.

(iii) Defn: $X_n \to X$ **in probability if** $\forall \epsilon > 0$,
$P(|X_n - X| > \epsilon) \to 0$ **as** $n \to \infty$.

(iv) Defn: $X_n \to X$ **in distribution, if the dfs** F_n **and** F **of** X_n **and** X **are s.t.** $F_n(x) \to F(x)$ **as** $n \to \infty$ **for each continuity point** x **of** F. **(**F **must be a proper df).**

(v) Let $0 < r < \infty$. $X_n \to X$ **in** r **th moment if**
$\mathrm{E}(|X_n - X|^r) \to 0$ **as** $n \to \infty$. **(**$\mathrm{E}(X_n^r)$ **and** $\mathrm{E}(X^r)$ **must** $\exists$**.)**

(vi) A sequence of rvs (X_n) **is uniformly integrable if** $\limsup_{n\to\infty} \mathrm{E}(|X_n| I(|X_n| \geq \lambda)) \to 0$ **as** $\lambda \to \infty$.

(vii) Notation

A: $X_n \to X$ **a.s.** **P:** $X_n \to X$ **in prob.**

D: $X_n \to X$ **in dsn.** **E$_r$:** $X_n \to X$ **in rth moment**

A$_S$: $X_{n'} \to_{a.s.} X$ **for subsequence** (n')

A$_{SS}$: Every subsequence (n') **contains** (n''), $X_{n''} \to X$ **a.s.**

UI$_r$: $|X_n|^r$ **are uniformly integrable.**

B$_r$: $|X_n| \leq Y$ **a.s.**$\forall n$ **with** $\mathrm{E}(Y^r) < \infty$

3.2 Convergence implications (JAW 2.4-8, Ferg Ch 1)

(i) A $\Rightarrow$ P (from the defn)

(ii) P $\Rightarrow$ D: let $A_n(\epsilon) \equiv \{|X_n - X| > \epsilon\}$
Let $\forall \epsilon > 0,\ \forall \delta > 0$ **s.t.** $P(A_n(\epsilon)) < \delta\ \forall\ n > n_0(\epsilon, \delta)$
$\{X_n \leq x\} \subseteq \{X \leq x + \epsilon\} \cup A_n(\epsilon)$, $\quad F_n(x) \leq F(x + \epsilon) + \delta$
$\{X_n > x\} \subseteq \{X > x - \epsilon\} \cup A_n(\epsilon)$, $\quad (1 - F_n(x)) \leq (1 - F(x - \epsilon)) + \delta$
So $F(x - \epsilon) - \delta \leq F_n(x) \leq F(x + \epsilon) + \delta$.
Taking x **a cty pnt of** F **and** $\epsilon \to 0$ **gives result.**

(iii) $\mathbf{A}_{SS} \Leftrightarrow$ P $\Rightarrow \mathbf{A}_S$
r.h.s: Choose n_k **st** $P(A_k \equiv \{|X_{n_k} - X| > 2^{-k}\}) < 2^{-k}$ **so** $P(\cup_{k \geq m} A_k) < 2 \times 2^{-m} \to 0$ **as** $m \to \infty$.
Now P $\Rightarrow \mathbf{A}_{SS}$ follows similarly. Conversely, if not P, then there is a subsequence not converging in probability, and hence not $\mathbf{A}_S$ (nor $\mathbf{A}_{SS}$ for any subsequence).

(iv) $\mathbf{E}_r \Rightarrow \mathbf{E}_s$ for $0 < s \leq r$
Holder's inequality gives $(\mathrm{E}(|X_n - X|^s))^{1/s} \leq (\mathrm{E}(|X_n - X|^r))^{1/r}$

(v) Vitali's Theorem (without proof):
If P and $\mathrm{E}(|X_n|^r) < \infty$, **then $\mathbf{UI}_r$ iff $\mathbf{E}_r$ iff** $\mathrm{E}(|X_n|^r) \to \mathrm{E}(|X|^r)$

(vi) Note: $\mathbf{UI}_r$ is a little weaker than B_r: example
$Z \sim U(0, 1)$, **let** $X_n = (n^\alpha / \log(1 + n)) I_{[0, 1/n^\alpha)}(Z)$. **Then if** $\alpha > 0$, $X_n \to_{a.s.} 0$, $E(X_n) \to 0$, **and $\mathbf{UI}_r$, for all** $r > 0$. **BUT $\mathbf{B}_r$ does not hold for any** $r > 0$.

(vii) $\mathbf{E}_r \Rightarrow$ P $\Rightarrow$ { $\mathbf{E}_r$ provided $\mathbf{B}_r$} (for any $r > 0$**)**
For l.h.s: $P(|X_n - X| > \epsilon) \leq \mathrm{E}(|X_n - X|^r)/\epsilon^r$.
For r.h.s., see (v), (vi).

3.3 Skorokhod theorems (JAW 2.13)

(i) Recall that any df F has at most a countable number of discontinuity points. Ditto F^{-1}.

(ii) For any df F, define $F^{-1}(u) = \inf\{x; F(x) \geq u\}$

(iii) If X has cts df F, $F(X) \sim U(0,1)$. For any F, $P(F(X) \leq t) \leq t$ with equality iff t is in the range of F.
(a) F & F^{-1} cts at t: $P(V \equiv F(X) \leq t) = F(F^{-1}(t)) = t$
(b) $F(x-\epsilon) = F(x) = t_0$: $F(F^{-1}(t)) = t$ as in (a),
(but note $F^{-1}(F(x)) = F^{-1}(t_0) \leq x - \epsilon < x$)
(c)$t \notin$ range(F): $F^{-1}(t) = x_0, \ F(x_0) > t$; i.e. $F(F^{-1}(t)) > t$

(iv) Inverse Transformation Theorem:
Let $V \sim U(0,1)$, $X = F^{-1}(V)$. Then X has df F.
(i.e. $V \leq F(x)$ iff $X \leq x$, so $P(X \leq x) = P(V \leq F(x)) = F(x)$.)
Proof: $V \leq F(x) \Rightarrow X \equiv F^{-1}(V) \leq x$ by defn F^{-1}
$F^{-1}(V) \equiv X \leq x \Rightarrow V \leq F(x+\epsilon), \forall \epsilon > 0$ i.e. $V \leq F(x)$.

(v) Skorokhod Theorem: Suppose $X_n \rightarrow_d X$, then $\exists$ X_n^*, X^* s.t. $X_n^* \rightarrow_{a.s.} X^*$, $X_n^* =_d X_n$, $X^* =_d X$.

Proof: Let $U \sim U(0,1)$, F_n be df of X_n, F of X
Let $X_n^* = F_n^{-1}(U)$, $X^* = F^{-1}(U)$. Clearly $X_n^* =_d X_n$, $X^* =_d X$.

Now, let t be a continuity point of F^{-1}, $z = F^{-1}(t)$, and x, y be cty pnts of F, $x < z < y$.
Then $F(x) < t$ for $x < z$, so $F_n^{-1}(t) \geq x$ for large enough n.
Also $F(y) > t$ for $z < y$, so $F_n^{-1}(t) \leq y$ for large enough n.
Now let $x \uparrow z, \ y \downarrow z$ through cty points of F.
So $F_n^{-1}(t) \rightarrow F^{-1}(t)$ except at a countable number of t, which has Lebesgue measure 0. i.e. $X_n^* \rightarrow_{a.s.} X^*$.
Note: Versions of Skorokhod hold also in $\Re^k-$ but multivariate monotonicity is not so simple.

3.4 Classical limit theorems (JAW 2.11, 14, Ferg 4, 5, 6)

(i) WLLN: X_i **i.i.d with mean** μ ($\mathrm{E}|X_1| < \infty$)**,** $\overline{X_n} \rightarrow_p \mu$**.**

(ii) SLLN: X_i **i.i.d with mean** μ ($\mathrm{E}|X_1| < \infty$)**,** $\overline{X_n} \rightarrow_{a.s} \mu$**.**

(iii) Multivariate CLT: X_i **i.i.d. in** $\Re^k$ **with mean** $\mathrm{E}(X_1) = \mu$ **and** $\mathrm{var}(X_1) = \Sigma = \mathrm{E}(X_1 - \mu)(X_1 - \mu)'$ **(so** $\mathrm{E}(X_1' X_1) < \infty$**), then** $n^{\frac{1}{2}}(\overline{X_n} - \mu) \rightarrow_d N_k(0, \Sigma)$

(iv) Other CLTs: Liapouvov and Lindeberg-Feller conditions – see JAW 2.11, Ferg, P27-28.

(v) Continuous mapping/Mann-Wald Theorem
Suppose $g : \Re^k \rightarrow \Re$ **is continuous a.s.(**P_X**) then**
$X_n \rightarrow_t X \;\Rightarrow\; g(X_n) \rightarrow_t g(X)$ **where** t **is** $a.s.$, p **or** d**.**

(vi) Cramér-Wold device: $X_n \in \Re^k$
$X_n \rightarrow_d X$ **iff** $a'X_n \rightarrow_d a'X \quad \forall a \in \Re^k$**.**

(vii) Helly-Bray Theorem: If $X_n \rightarrow_d X$ **and** g **bounded and cts a.s.** P_X **then** $\mathrm{E}(g(X_n)) \rightarrow \mathrm{E}(g(X))$ **as** $n \rightarrow \infty$**.**

(viii) Slutsky's Thm: simple case
$A_n \rightarrow_p a$**,** $B_n \rightarrow_p b$**,** $Z_n \rightarrow_d X \quad \Rightarrow \quad (A_n Z_n + B_n) \rightarrow_d (aX + b)$

(ix) Slutsky corollary: simple g' **Theorem (**δ**-method)**
If X **is a proper r.v. and** g **is ctsly diffble at** b**,** $a_n \rightarrow \infty$
$a_n(Z_n - b) \rightarrow_d X \quad \Rightarrow \quad a_n(g(Z_n) - g(b)) \rightarrow_d g'(b)X$

(x) Slutsky's Thm (multiparameter version)
$g : \Re^k \rightarrow \Re$**,** $b \in \Re^k$**,** $\bigtriangledown g = (\frac{\partial g}{\partial b_i})$ $\exists$ **cts.** $Z_n,\ X$ **are r.vs in** $\Re^k$**,**
$a_n \rightarrow \infty$**,** $\quad a_n(Z_n - b) \rightarrow_d X \quad \Rightarrow \quad a_n(g(Z_n) - g(b)) \rightarrow_d (\bigtriangledown g)'X$

(xi) Corollary: If g **has cts partial deriv at** b**,** $a_n \rightarrow \infty$ **and** $a_n(Z_n - b) \rightarrow_d X \sim N_k(0, \Sigma)$ **where** $a_n \rightarrow \infty$ **then** $a_n(g(Z_n) - g(b)) \rightarrow_d (\bigtriangledown g)'X \sim N(0, (\bigtriangledown g)'\Sigma(\bigtriangledown g))$

3.5 Empirical processes, Brownian Bridge (JAW 2.16-18)

(i) Let U_i, $i = 1, 2, ...$ be i.i.d. $U(0,1)$. Define their empirical dsn fn (edf) $G_n(t) = n^{-1} \sum_{i=1}^{n} I\{U_i \leq t\}$.
Note $nG_n(t) \sim Bin(n,t)$: $\mathrm{E}(G_n(t)) = t,\ \mathrm{var}(G_n(t)) = t(1-t)/n$.

(ii)The uniform empirical process is $U_n(t) \equiv n^{\frac{1}{2}}(G_n(t) - t)$.

(iii) $\mathrm{E}(U_n(t)) = 0$**,** $\mathrm{var}(U_n(t)) = t(1-t)$**, and** $\mathrm{Cov}(U_n(t), U_n(s)) = n\mathrm{Cov}(G_n(t), G_n(s)) = \mathrm{Cov}(I\{U_i \leq s\} I\{U_i \leq t\}) = \min(s,t) - st$**.**

(iv) A Gaussian process $B(t),\ 0 < t < 1$, with this mean and covariance is Brownian Bridge.
By the multivariate CLT
$(U_n(t_1), ..., U_n(t_k)) \rightarrow_d (B(t_1), ..., B(t_k)) \sim N_k(0, \Sigma)$
where $\sigma_{ij} = \min(t_i, t_j) - t_i t_j$, for any finite k.

(v) Let X_i, $i = 1, 2, ...$ be i.i.d. with cdf F. Their edf is $F_n(x) = n^{-1} \sum_{i=1}^{n} I\{X_i \leq x\}$**, and the corresponding empirical process is** $n^{\frac{1}{2}}(F_n(x) - F(x))$**.**

(vi) Consider $X_i^* = F^{-1}(U_i) \sim F$, and $X_i^* \leq x$ iff $U_i \leq F(x)$ a.s. So $F_n^*(x) = G_n(F(x))$, $-\infty < x < \infty$.

(vii) Then $n^{\frac{1}{2}}(F_n - F) \ =_d\ n^{\frac{1}{2}}(F_n^* - F) \ =_d$

$$n^{\frac{1}{2}}(G_n(F) - F) \ \equiv\ U_n(F) \ \rightarrow_d\ B(F)$$

where cgce is of the finite-dimensional dsns.

(viii) Glivenko-Cantelli Theorem: for any F,
$\sup_x |F_n(x) - F(x)| \rightarrow_{a.s.} 0$ **as** $n \rightarrow \infty$**.**
Proof: $\sup_x |F_n(x) - F(x)| =_d \sup_x |F_n^*(x) - F(x)| = \sup_x |G_n(F(x)) - F(x)| \leq \sup_t |G_n(t) - t|$**. So suffices to show for G_n and $0 < t < 1$. Now split $(0,1)$ into bits $I_j = ((j-1)\delta, j\delta)$ size $\delta = M^{-1}$ for large M, and look at the** $\max_j \sup_{t \in I_j} |G_n(t) - t|$**, and use $G_n(t)$ incr, and $G_n(j\delta) \rightarrow_{a.s.} j\delta$.**

3.6 Limit theory for Uniform sample quantiles (JAW 2.28-30)

(i) Let $U_{(i)}$ $i = 1, ..., n$ **be order statistic of** n**-sample of** $U(0,1)$ $G_n^{-1}(t) = \inf\{u : G_n(u) \geq t\} = U_{(i)}$ **for** $(i-1)/n < t \leq i/n$**.**

(ii) The uniform quantile function is G_n^{-1}**, and the uniform quantile process is** $Q_n(t) = n^{\frac{1}{2}}(G_n^{-1}(t) - t)$**.**

(iii) $\sup_{0<t<1} |G_n^{-1}(t) - t| \;\leq\; \sup_{0<t<1} |G_n(t) - t| + (1/n)$ **(picture) which** $\to_{a.s.} 0$ **by Glivenko-Cantelli.**

(iv) $n^{\frac{1}{2}}(U_{(i)} - p) \;\to_d N(0, p(1-p))$ **if** $n^{\frac{1}{2}}((i/n) - p) \to 0$**.**
(a) $U_{(i)} \leq w$ **iff** $(\# U_j \leq w) \geq i$ **iff** $B(n,w) \geq i$ **prob (approx)**

$$
\begin{aligned}
P(N(nw, nw(1-w)) \geq i) &= \Phi((nw - i)/\sqrt{nw(1-w)}) \\
P(n^{\frac{1}{2}}(U_{(i)} - p) \leq x) &= P(U_{(i)} \leq x n^{-\frac{1}{2}} + p) \;\approx \\
\Phi\left(\frac{n^{\frac{1}{2}}x + np - i}{\sqrt{np(1-p) + O(n^{\frac{1}{2}})}}\right) &\approx \Phi(-x/\sqrt{p(1-p)})
\end{aligned}
$$

This proof can be made rigorous.
(b) $(U_{(1)}, ..., U_{(n)}) \;=_d\; (W_j/W_{n+1}; j = 1, ..., n)$ **where** $W_j = \Sigma_{k=1}^{j} V_k$**, where** $V_1, ..., V_{n+1}$ **are i.i.d. exponentials mean 1. (See Hwk 5/6 ?). This gives easy proof for each** $U_{(j)}$**, and a neater proof for joint dsn of the** $(U_{(j)})$**.**
(c) Using empirical processes – see JAW notes.

(v) For the finite-dimensional dns

$$
(n^{\frac{1}{2}}(G_n^{-1}(t_j) - t_j); j = 1, ..., k) \;\to_d\; N_k(0, \Sigma)
$$

where $\sigma_{ij} = \min(t_i, t_j) - t_i t_j$**.**
Or, if $Q_n(t) = n^{\frac{1}{2}}(G_n^{-1}(t) - t)$**,** $\sup_t |Q_n(t) - B(t)| \to_{a.s.} 0$ **where** $B()$ **is the Brownian Bridge process.**

3.7 Quantiles of a general dsn (Ferg 13)

(i) $F_n(x) = n^{-1}\Sigma_{i=1}^n I\{X_i \le x\}$**, where** $X_1, ..., X_n$ **are i.i.d.** $\sim F$**.** $F_n(x) \to_{a.s.} \mathrm{E}(I\{X_i \le x\}) \;=\; F(x)$ **by SLLN (see 3.5 (v)).**

(ii) $F_n^{-1}(t) \equiv \inf\{x : F_n(x) \ge t\} = X_{(i)}$ **on** $(i-1)/n < t \le i/n$**. By inverse transformation thm**

$$(F^{-1}(U_{(1)}), ..., F^{-1}(U_{(n)})) \;=_d\; (X_{(1)}, ..., X_{(n)})$$

or $F^{-1}(G_n^{-1}()) =_d F_n^{-1}()$**, at least for finite dim dsns.**

(iii) Now from 3.6 (iii),
$\sup_{a\le t\le b}|G_n^{-1}(t) - t| \;\le\; \sup_{0\le t\le 1}|G_n^{-1}(t) - t| \;\to_{a.s.} 0$**, so if** F^{-1} **is continuous on** $[a, b]$

$$\sup_{a\le t\le b}|F_n^{-1}(t) - F^{-1}(t)| \;=_d\; \sup_{a\le t\le b}|F^{-1}(G_n^{-1}(t)) - F^{-1}(t)| \;\to_{a.s.} 0$$

(iv) From 3.6 (iv), $n^{\frac{1}{2}}(G_n^{-1}(p) - p) \;\to_d\; N(0, p(1-p))$**,**
So, if F **is absolutely continuous, with pdf** $f = F'$**, and** $f(F^{-1}(p)) > 0$

$$n^{\frac{1}{2}}(F_n^{-1}(p) - F^{-1}(p)) \;\to_d\; Q'(p).N(0, p(1-p))$$

where $Q = F^{-1}$**, so** $Q'(p) = 1/f(F^{-1}(p))$
"Proof": linearize as for g' **theorem, or** δ**-method.**

(v) In general, for the finite-dimensional dns of Brownian Bridge process $B(t)$**:**

$$(n^{\frac{1}{2}}(F_n^{-1}(t_j) - F^{-1}(t_j)), j = 1, ..., k) \;\to_d\; (Q'(t_j)B(t_j); j = 1, ..., k)$$

"Proof": linearize as for multivariate g' **theorem, or** δ**-method.**