CHAPTER 2.

SOME BASIC LARGE SAMPLE THEORY

1. **Modes of Convergence**
2. **Classical Limit Theorems**
3. **Replacing $\rightarrow_d$ by $\rightarrow_{a.s.}$; Skorokhod's theorem**
4. **Empirical Distribution Functions, Empirical Measures, and Empirical Processes and Brownian bridge**
   The uniform empirical process
   The Brownian bridge process
   Finite - dimensional convergence
   A hint of Donsker's Mann - Wald type theorem
   The empirical process on R .
   General empirical measures and processes
5. **The Partial Sum Process and Brownian Motion**
   Brownian motion
   The partial sum process
   Existence of Brownian Motion and Brownian Bridge as Continuous Processes on $C[0,1]$.
   Finite - dimensional convergence
6. **Sample quantiles**

1. Modes of Convergence

Consider a probability space $(\Omega, \mathbf{A}, P)$. For our first two definitions we suppose that X , X_n , $n \geq 1$ are all random variables defined on this one probability space.

Definition 1.1. We say that X_n *converges a.s.* to X , denoted by $X_n \rightarrow_{a.s.} X$, if

(1) $$X_n(\omega) \rightarrow X(\omega) \quad \text{for all} \quad \omega \in A \quad \text{where} \quad P(A^c) = 0 ,$$

or, equivalently, if, for every $\epsilon > 0$

(2) $$P(\sup_{m \geq n} |X_m - X| > \epsilon) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty .$$

Definition 1.2. We say that X_n *converges in probability to* X and write $X_n \rightarrow_p X$) if for every $\epsilon > 0$

(3) $$P(|X_n - X| \geq \epsilon) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty .$$

Definition 1.3. Let $0 < r < \infty$. We say that X_n *converges in* $r - th$ *mean* to X , denoted by $X_n \rightarrow_r X$, if

(4) $$E|X_n - X|^r \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty \quad \text{for functions} \quad X_n, X \quad \text{in} \quad \mathbf{L}_r .$$

Definition 1.4. We say that X_n *converges in distribution* to X (denoted by $X_n \rightarrow_d X$, $F_n \rightarrow F$ or $\mathbf{L}(X_n) \rightarrow \mathbf{L}(X)$, with $\mathbf{L}$ referring to "law") if the df's F and F_n of X and X_n satisfy

(5) $$F_n(x) \rightarrow F(x) \quad \text{as} \quad n \rightarrow \infty \quad \text{for each continuity point} \quad x \quad \text{of} \quad F .$$

(Note that $F_n \equiv 1_{[1/n, \infty)} \rightarrow_d F \equiv 1_{[0,\infty)}$ even though $F_n(0) = 0$ *does not converge* to $1 = F(0)$.) The statement $\rightarrow_d$ will carry with it the implication that F corresponds to a (proper) probability measure P .

Definition 1.5. A sequence of random variables $\{X_n\}$ is *uniformly integrable* if

$$\lim_{\lambda \rightarrow \infty} \operatorname{limsup}_{n \rightarrow \infty} E|X_n| 1_{[|X_n| \geq \lambda]} = 0 .$$

Theorem 1.6. (Convergence implications).

A. If $X_n \rightarrow_{a.s.} X$, then $X_n \rightarrow_p X$.

B. If $X_n \rightarrow_p X$, then $X_{n'} \rightarrow_{a.s.} X$ for some subsequence n' .

C. If $X_n \rightarrow_r X$, then $X_n \rightarrow_p X$.

D. If $X_n \to_p X$ and $|X_n|^r$ is uniformly integrable, then $X_n \to_r X$.
If $X_n \to_p X$ and $\overline{\lim} E|X_n|^r \le E|X|^r < \infty$, then $X_n \to_r X$.

E. If $X_n \to_r X$ then $X_n \to_{r'} X$ for all $0 < r' \le r$.

F. If $X_n \to_p X$, then $X_n \to_d X$.

G. $X_n \to_p X$ if and only if every subsequence $\{n'\}$ contains a further subsequence $\{n''\}$ for which $X_{n''} \to_{a.s.} X$.

Theorem 1.7. (Vitali's theorem). Suppose $X_n \in L_r(P)$ where $0 < r < \infty$ and $X_n \to_p X$. Then the following are equivalent:

A. $|X_n|^r$ are uniformly integrable.

B. $X_n \to_r X$

C. $E|X_n|^r \to E|X|^r$.

Before proving the theorems we need a short review of some facts about convex functions and some inequalities. We first briefly review convexity. A real valued function f is *convex* if

(6) $$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$$

for all x, y and all $0 \le \alpha \le 1$. This holds if and only if

(7) $$f(\frac{1}{2}(x + y)) \le \frac{1}{2}[f(x) + f(y)]$$

for all x, y provided f is continuous and bounded. Note that (7) holds if and only if

(8) $$f(s) \le \frac{1}{2}[f(s-r) + f(s+r)] \quad \text{for all } r, s.$$

Also

(9) $$f''(x) \ge 0 \quad \text{for all } x \text{ implies } f \text{ is convex.}$$

We call f *strictly convex* if strict inequality inequality holds in any of the above. If f is convex, then there exists a linear function l such that $f(x) \ge l(x)$ with equality at a prespecified x_0 in the interior of the domain of f; this is called the *supporting hyperplane theorem.*

Definition 1.8. Assuming the following expectations (integrals) exist,

(10) $\mu \equiv EX =$ the *mean* of X.

(11) $\sigma^2 \equiv Var[X] \equiv E(X - \mu)^2 =$ the *variance* of X.

(12) $EX^k =$ k th *moment* of X for $k \ge 1$ an integer.

(13) $E|X|^r = r$*- th absolute moment* of X for $r \ge 0$.

(14) $E(X - \mu)^k = k$ - *th central moment* of X .

(15) $Cov[X,Y] \equiv E[(X - \mu_X)(Y - \mu_Y)] =$ the *covariance* of X and Y .

Proposition 1.9. If $E|X|^r < \infty$, then $E|X|^{r'}$ and EX^k are finite for all $r', k \le r$.

Proof. Now $|x|^{r'} \le 1 + |x|^r$; and integrability is equivalent to absolute integrability. □

Proposition 1.10. $Var(X) \equiv \sigma^2 < \infty$ if and only if $EX^2 < \infty$. In this case $\sigma^2 = EX^2 - \mu^2$.

Proof. Suppose that $\sigma^2 < \infty$. Then $\sigma^2 + \mu^2 = E(X - \mu)^2 + E(2\mu X - \mu^2) = EX^2$.
Suppose that $EX^2 < \infty$. Then $EX^2 - \mu^2 = E(X^2) - E(2\mu X - \mu^2) = E(X - \mu)^2 = Var[X]$. □

Inequality 1.11. ($c_r -$ inequality). $E|X + Y|^r \le c_r E|X|^r + c_r E|Y|^r$ where $c_r = 1$ for $0 < r \le 1$ and $c_r = 2^{r-1}$ for $r \ge 1$.

Proof. Case 1: $r \ge 1$. Then $|x|^r$ is a convex function of x ; take second derivatives. Thus $|(x + y)/2|^r \le [\,|x|^r + |y|^r\,]/2$; and now take expectations.
Case 2: $0 < r \le 1$: Now $|x|^r$ is concave and ↑ for $x \ge 0$; examine derivatives. Thus

$$\begin{aligned} |x + y|^r - |x|^r &= \int_x^{x+y} rt^{r-1}\,dt = \int_0^y r(x + s)^{r-1}\,ds \\ &\le \int_0^y r(s)^{r-1}\,ds = |y|^r , \end{aligned}$$

and now take expectations. □

Inequality 1.12. (Hölder inequality). $E|XY| \le E^{1/r}|X|^r E^{1/s}|Y|^s$ for $r > 1$ where $\frac{1}{r} + \frac{1}{s} = 1$ defines s . When the expectations are finite we have equality if and only if there exists A and B not both 0 such that $A\,|X|^r = B\,|Y|^s$ a.e..

Proof. The result is trivial if $E|X|^r = 0$ or ∞ . Likewise for $E|Y|^s$. So suppose > 0 . Now

$$|ab| \le \frac{|a|^r}{r} + \frac{|b|^s}{s}, \qquad \text{as in the figure .}$$

Now let $a = |X|/\,||X||_r$ and $b = |Y|/\,||Y||_s$; and take expectations. Equality holds if and only if $|Y|/||Y||_s = (|X|/||X||_r)^{1/(1-s)}$ a.e.; if and only if

$$\frac{|Y|^s}{E|Y|^s} = (\frac{|X|}{||X||_r})^{\frac{s}{s-1}} = \frac{|X|^r}{E|X|^r} \quad \text{a.e.}$$

if and only if there exist $A, B \neq 0$ such that $A\,|X|^r = B\,|Y|^s$. This also gives the next inequality as an immediate corollary. □

Inequality 1.13. (Cauchy - Schwarz inequality). $(E|XY|)^2 \leq EX^2\,EY^2$ with equality if and only if there exists A, B not both 0 such $A\,|X| = B\,|Y|$ a.e..

Remark Thus for non-degenerate r.v.'s (i.e. non - zero variance) with finite variance we have

(16) $$-1 \leq \rho \leq 1$$

where

(17) $$\rho \equiv Corr\,[X,Y] \equiv \frac{Cov[X,Y]}{\sqrt{Var[X]\,Var[Y]}} .$$

is called the *correlation* of X and Y . Note that $\rho = 1$ if and only if $X - \mu_X = A(Y - \mu_Y)$ for some $A > 0$ and $\rho = -1$ if and only if $X - \mu_X = A(Y - \mu_Y)$ for some $A < 0$. Thus ρ measures linear dependence, not dependence.

Proposition 1.14. $\log E|X|^r$ is convex in r for $r \geq 0$. It is linear if and only if $|X| = c$ a.e. for some c .

Proof. Let $0 \leq r \leq s$. Apply the Cauchy-Schwartz inequality to $|X|^{(s-r)/2}$ and $|X|^{(s+r)/2}$ and take logs to get

$$\log E|X|^s \leq \frac{1}{2}\,[\log E|X|^{s-r} + \log E|X|^{s+r}] .$$

□

Inequality 1.15. (Liapounov Inequality). Let X be a random variable. Then $E^{1/r}|X|^r$ is $\uparrow$ in r for $r \geq 0$.

Proof. The slope of the chord of $y = \log E|X|^r$ is $\uparrow$ in r by proposition 1.14. That is, $\frac{1}{r}\log E|X|^r$ is $\uparrow$ in r . We used $P(\Omega) = 1 < \infty$. □

Exercise 1.16. Let $\mu_r \equiv E|X|^r$. For $r \geq s \geq t \geq 0$ we have $\mu_r^{s-t}\,\mu_t^{r-s} \geq \mu_s^{r-t}$.

Inequality 1.17. (Minkowski's inequality). For $r \geq 1$ we have $E^{1/r}|X + Y|^r \leq E^{1/r}|X|^r + E^{1/r}|Y|^r$.

Proof. It is trivial for $r = 1$. Suppose that $r > 1$. Then for any measure μ

$$\text{(a)} \qquad E|X + Y|^r \le E|X||X+Y|^{r-1} + E|Y||X+Y|^{r-1}$$
$$\le (||X||_r + ||Y||_r)\,|||X + Y|^{r-1}||_s \qquad \text{by Hölder's inequality}$$
$$= (||X||_r + ||Y||_r)\,E^{1/s}\,|X + Y|^{(r-1)s}$$
$$= (||X||_r + ||Y||_r)\,E^{1/s}\,|X + Y|^{r}\,.$$

If $E|X + Y|^r = 0$, it is trivial. If not, we divide to get the result. □

Inequality 1.18. (Basic inequality). Let $g \ge 0$ be an even function which is $\uparrow$ on $[0,\infty)$. Then for all rv's X and for all $\epsilon > 0$

$$P(\,|X| \ge \epsilon\,) \le \frac{Eg(X)}{g(\epsilon)}\,.$$

Proof. Now

$$\text{(a)} \qquad Eg(X) = \int_{|X| \ge \epsilon} g(X)\,dP + \int_{|X| < \epsilon} g(X)\,dP \ge \int_{|X| \ge \epsilon} g(X)\,dP$$
$$\ge g(\epsilon)\int_{|X| \ge \epsilon} 1\,dP = g(\epsilon)\,P(\,|X| \ge \epsilon\,)$$

as claimed. □

The next two inequalities are immediate corollaries.

Inequality 1.19. (Markov's inequality). $P(\,|X| \ge \epsilon\,) \le \dfrac{E|X|^r}{\epsilon^r}$ for all $\epsilon > 0$.

Inequality 1.20. (Chebyshev's inequality). $P(\,|X - \mu| \ge \epsilon\,) \le \dfrac{Var[X]}{\epsilon^2}$ for all $\epsilon > 0$.

Inequality 1.21. (Jensen's inequality). If g is convex on (a,b) where $-\infty \le a < b \le \infty$ and if $P(X \in (a,b)) = 1$ and EX is finite (and hence $a < EX < b$), then

$$\text{(18)} \qquad g(\,EX\,) \le Eg(X)\,.$$

For strict convexity we have equality if and only if $X = EX$ a.s.

Proof. Let l be a supporting hyperplane to g at EX. Then

$$\text{(a)} \qquad Eg(X) \ge El(X)$$
$$\text{(b)} \qquad = l(EX) \qquad \text{since } l \text{ is linear and } P(\Omega) = 1$$
$$\text{(c)} \qquad = g(EX)\,.$$

Now $g(X) - l(X) \ge 0$. Thus $Eg(X) = El(X)$ if and only if $g(X) = l(X)$ a.s. if and only if $X = EX$ a.s.. □

Exercise 1.22. For any function $h \in \mathbf{L}_2(0,1)$, define a new function Th on $(0,1)$ by $Th(u) = u^{-1}\int_0^u h(s)\,ds$ for $0 < u \le 1$. Use the Cauchy - Schwarz inequality to show that

$$\int_0^1 \{Th(u)\}^2\,du \le 4\int_0^1 h^2(u)\,du\ .$$

Thus $T : L_2(0,1) \to L_2(0,1)$ is a *bounded linear operator* with $||T|| \le 2$. [Hint: write $Th(u) = u^{-1}\int_0^u h(s)\,s^{\alpha}\,s^{-\alpha}\,ds$ for some α .]

Exercise 1.23. Suppose that $X \sim Binomial(n\,,p)$. Use the basic inequality 1.18 with $g(x) = \exp(rx)$, $r > 0$, to show that for $\epsilon \ge 1$

$$P(\frac{X/n}{p} \ge \epsilon) \le \exp(-\,np\,h(\epsilon))$$

where $h(\epsilon) \equiv \epsilon(\log(\epsilon) - 1) + 1$. From this, show that for $\lambda > 0$ we have

$$P(\sqrt{n}(\frac{X}{n} - p) \ge \lambda) \le \exp(-\frac{\lambda^2}{2p}\,\psi(\frac{\lambda}{p\sqrt{n}}))$$

where $\psi(x) \equiv 2h(1+x)/x^2$ is monotone decreasing on $[0,\infty)$ with $\psi(0) = 0$ and $\psi(x) \sim 2\log(x)/x$ as $x \to \infty$.

Proof of theorem 1.6. A follows easily since

$$P(|X_n - X| \ge \epsilon) \le P(\bigcup_{m \ge n}[|X_m - X| \ge \epsilon]) \to 0\ .$$

To prove B, first note that $X_n \to_p X$ implies that for every $k \ge 1$ there exists an integer n_k such that $P(|X_{n_k} - X| > 1/2^k) < 2^{-k}$; we can assume $n_k \uparrow$ in k; if not take $n'_k \equiv n_k + k$. Let $A_m \equiv \cup_{k \ge m}[|X_{n_k} - X| > 2^{-k}]$ so that $P(A_m) \le \sum_{k=m}^{\infty} 2^{-k} = 2^{-m+1}$. On $A_m^c = \cap_{k \ge m}[|X_{n_k} - X| \le 2^{-k}]$, $|X_{n_k} - X| \le 2^{-k}$ for all $k \ge m$; i.e on A_m^c, $X_{n_k}(\omega) \to X(\omega)$. Thus $X_{n_k} \to X$ on $A \equiv \cup_{m=1}^{\infty} A_m^c$, and $P(A^c) = P(\cap_{m=1}^{\infty} A_m) = \lim_m P(A_m) \le \lim_m 2^{-m+1} = 0$.

Markov's inequality gives C via $P(|X_n - X| \ge \epsilon) \le E|X_n - X|^r/\epsilon^r \to 0$. Hölder's inequality with $\frac{1}{r/r'} + \frac{1}{q} = 1$ gives E via

$$E|X_n - X|^{r'} \le \{E|X_n - X|^{r'(r/r')}\}^{r'/r}\{E1^q\}^{1/q}$$

$$\text{(a)} \qquad = \{E|X_n - X|^r\}^{r'/r} \to 0\,;$$

or use Liapunov's inequality.

Vitali's theorem gives D.

Consider F. Let $X_n \sim F_n$ and $X \sim F$. Now

(b) $$F_n(t) = P(X_n \le t) \le P(X \le t+\epsilon) + P(|X_n - X| \ge \epsilon)$$

(c) $$\le F(t+\epsilon) + \epsilon \quad \text{for all } n \ge \text{some } n_\epsilon .$$

Also

$$\begin{aligned} F_n(t) &= P(X_n \le t) \ge P(X \le t-\epsilon \text{ and } |X_n - X| \le \epsilon) \equiv P(AB) \\ &\ge P(A) - P(B^c) = F(t-\epsilon) - P(|X_n - X| \ge \epsilon) \\ &\ge F(t-\epsilon) - \epsilon \quad \text{for } n \ge \text{some } n'_\epsilon . \end{aligned}$$

Thus

(e) $$F(t-\epsilon) - \epsilon \le \underline{\lim}\, F_n(t) \le \overline{\lim}\, F_n(t) \le F(t+\epsilon) + \epsilon .$$

If t is a continuity point of F, then letting $\epsilon \to 0$ in (e) gives $F_n(t) \to F(t)$. Thus $F_n \to_d F$.

Half of G follows from B since any $X_{n'} \to_p X$. We turn to the other half. Assume that $X_n \to_p X$ fails. Then there exists $\epsilon_0 > 0$ for which $\delta_0 \equiv \overline{\lim}\, P(|X_n - X| \ge \epsilon_0) > 0$. Thus there exists a subsequence $\{n'\}$ for which $P(|X_{n'} - X| \ge \epsilon_0) \to \delta_0 > 0$. Thus neither $X_{n'}$ nor any further subsequence can $\to_p X$. Thus neither $X_{n'}$ nor any further subsequence $X_{n''}$ can $\to_{a.s.} X$. This is a contradiction. Thus $X_n \to_p X$. □

Proof of theorem 1.7. Postponed.

Exercise 1.24. (Scheffé's theorem). If $f, f_1, f_2, \cdots$ are densities with respect to a measure μ and $f_n \to f$ a.e. μ, then with $P_n(B) \equiv \int_B f_n d\mu$ and $P(B) \equiv \int_B f d\mu$ for any Borel set B,

$$\sup_{A \in \mathbf{B}_n} |P_n(A) - P(A)| = \sup_A \left| \int_A f_n\, d\mu - \int_A f\, d\mu \right| = \frac{1}{2} \int |f_n - f|\, d\mu \to 0$$

as $n \to \infty$.

2. Classical Limit Theorems.

We now state some of the classical limit theorems of probability theory which are of frequent use in statistics.

Proposition 2.1. (WLLN). If $X_1, \cdots, X_n, \cdots$ are iid with mean μ (so $E|X| < \infty$ and $\mu = E(X)$), then $\bar{X}_n \to_p \mu$.

Proposition 2.2. (SLLN). If $X_1, \cdots, X_n, \cdots$ are iid with mean μ (so $E|X| < \infty$ and $\mu = E(X)$), then $\bar{X}_n \to_{a.s.} \mu$.

Proposition 2.3. (CLT). If $X_1, \cdots, X_n$ are iid (μ, σ^2) (so $E|X|^2 < \infty$, then $\sqrt{n}(\bar{X}_n - \mu) \to_d N(0, \sigma^2)$.

Proposition 2.4. (Multivariate CLT). If $X_1, X_2, \cdots, X_n$ are iid random vectors in R^d with $(\mu, \Sigma) \equiv (EX, E(X-\mu)(X-\mu)^T)$ (so $E(X^T X) < \infty$), then $\sqrt{n}(\bar{X}_n - \mu) \to_d N_d(0, \Sigma)$.

Proposition 2.5. (Liapunov CLT). Let $X_{n1}, \cdots, X_{nn}$ be row independent rv's with $\mu_{ni} \equiv EX_{ni}$, $\sigma_{ni}^2 \equiv Var(X_{ni})$, and $\gamma_{ni} \equiv E|X_{ni} - \mu_{ni}|^3 < \infty$. Let $\mu_n \equiv \sum_{i=1}^n \mu_{ni}$, $\sigma_n^2 = \sum_{i=1}^n \sigma_{ni}^2$, $\gamma_n \equiv \sum_{i=1}^n \gamma_{ni}$. If $\gamma_n / \sigma^3 \to 0$, then $(\sum_i X_{ni} - \mu_n)/\sigma_n \to_d N(0,1)$.

Proposition 2.6. (Lindeberg - Feller CLT). Let $X_{ni} \sim (0, \sigma_{ni}^2)$ be row independent. Let $S_n \equiv \sum_{i=1}^n X_{ni}$, $\sigma_n^2 = \sum_i \sigma_{ni}^2$. Then both $S_n/\sigma_n \to_d N(0,1)$ and $\max\{\sigma_{ni}^2/\sigma_n^2 : 1 \le i \le n\} \to 0$ if and only if the Lindeberg condition

$$\frac{1}{\sigma_n^2} \sum_{i=1}^n E|X_{ni}|^2 1_{[|X_{ni}| \ge \epsilon\sigma_n]} \to 0 \quad \text{for all} \quad \epsilon > 0 \quad \text{holds}.$$

Proposition 2.7. (The Cramér - Wold device). Random vectors $\underline{X}_n$ in R^d satisfy $\underline{X}_n \to_d \underline{X}$ if and only if $\underline{a}^T \underline{X}_n \to_d \underline{a}^T \underline{X}$ in R for all $\underline{a} \in R^d$.

Proposition 2.8. (Continuous mapping, or Mann - Wald theorem). Suppose that $g : R^d \to R$ is continuous a.s. P_X. Then:

A. If $X_n \to_{a.s.} X$, then $g(X_n) \to_{a.s.} g(X)$.

B. If $X_n \to_p X$, then $g(X_n) \to_p g(X)$.

C. If $X_n \to_d X$, then $g(X_n) \to_d g(X)$.

Proposition 2.9. (Slutsky's theorem). Suppose that $A_n \to_p a$, $B_n \to b$, where a, b are constants, and $X_n \to_d X$. Then $A_n X_n + B_n \to_d aX + b$.

Proposition 2.10. ($g'-$ theorem --or-- the delta method). Suppose that $a_n(X_n - b) \to_d X$ in R^m where $a_n \to \infty$, and suppose that

$g : R^m \rightarrow R^k$ has a derivative g' at b ; here g' is a $k \times m$ matrix. Then

$$a_n(g(X_n) - g(b)) \rightarrow_d g'(b)X \ .$$

Proposition 2.11. (Polya - Cantelli lemma). If $F_n \rightarrow_d F$ and F is continuous, then $||F_n - F||_\infty \rightarrow 0$.

3. Replacing $\to_d$ by $\to_{a.s.}$

Our goal in this section is to show how we can convert convergence in distribution into the stronger mode of almost sure convegence. This often simplifies proofs and makes them more intuitive.

Definition 3.1. For any df F define F^{-1} by $F^{-1}(t) \equiv \inf\{x : F(x) \geq t\}$ for $0 < t < 1$.

Proposition 3.2. F^{-1} is left - continuous.

Proof. To show that F^{-1} is left - continuous, let $0 < t < 1$, and set $z \equiv F^{-1}(t)$. Then $F(z) \geq t$ by the right continuity of F. If F is discontinuous at z, $F^{-1}(t-\epsilon) = z$ for all small $\epsilon > 0$, and hence left continuity holds. If F is continuous at z, then assume F^{-1} discontinuous from the left at t. Then for all $\epsilon > 0$, $F^{-1}(t-\epsilon) < z - \delta$ for some $\delta > 0$, and hence $F(z-\delta) \geq t - \epsilon$ for all $\epsilon > 0$. Hence $F(z-\delta) \geq t$, which implies $F^{-1}(t) \leq z - \delta$, a contradiction.

Proposition 3.3. If X has continuous df F, then $F(X) \sim \mathit{Uniform}(0,1)$. For *any* df F and any $0 < t < 1$,

$$P(F(X) \leq t) \leq t$$

with equality iff t is in the range of F. Equivalently, $F \circ F^{-1}(t) \geq t$ for all $0 < t < 1$ with equality iff t is in the range of F. Also, $F^{-1} \circ F(x) \leq x$ for all $-\infty < x < \infty$ with strict inequality iff $F(x-\epsilon) = F(x)$ for some $\epsilon > 0$. Thus $P(F^{-1} \circ F(X) \neq X) = 0$ where $X \sim F$.

Proof. Homework.

Theorem 3.4. (The inverse transformation.) Let $\xi \sim \mathrm{Uniform}(0,1)$ and let $X = F^{-1}(\xi)$. Then for all real x,

$$[X \leq x] = [\xi \leq F(x)] .$$

Thus X has distribution function F.

Proof. Now $\xi \leq F(x)$ implies $X = F^{-1}(\xi) \leq x$ by the definition 3.1 of F^{-1}. If $X = F^{-1}(\xi) \leq x$, then $F(x+\epsilon) \geq \xi$ for all $\epsilon > 0$, so that right continuity of F implies $F(x) \geq \xi$. We have thus shown that the asserted event identity holds. $\square$

Proposition 3.5. (Elementary Skorokhod theorem). Suppose that $X_n \to_d X_0$. Then there exist rv's X_n^*, $n \geq 0$, defined on $([0,1], \mathbf{B}[0,1], \lambda)$ for which $X_n^* \stackrel{d}{=} X_n$ for every $n \geq 0$ and $X_n^* \to_{a.s.} X_0^*$.

Proof. Let F_n denote the df of X_n and let

(a) $$X_n^* \equiv F_n^{-1}(\xi) \qquad \text{for all} \quad n \geq 0$$

where $\xi \sim Uniform(0,1)$. Then $X_n^* =_d X_n$ for all $n \geq 0$ by theorem 5.4. It remains only to show that $X_n^* \to_{a.s.} X_0^*$

Let $t \in (0,1)$ be such that there is at most one value z having $F(z) = t$.
Let $z = F^{-1}(t)$. Then $F(x) < t$ for $x < z$.
Thus $F_n(x) < t$ for $n \geq$ some N_x provided $x < z$ is a continuity point of F.
Thus $F_n^{-1}(t) \geq x$ for $n \geq N_x$ provided $x < z$ is a continuity point of F.
Thus $\underline{\lim}\, F_n^{-1}(t) \geq x$ provided $x < z$ is a continuity point of F.
Thus $\underline{\lim}\, F_n^{-1}(t) \geq z$ since there are continuity points x that $\uparrow z$.

We also have $F(x) > t$ for $z < x$.
Thus $F_n(x) > t$, and hence $F_n^{-1}(t) \leq x$ for $n \geq$ some N_x provided $x > z$ is a continuity point of F.
Thus $\overline{\lim}\, F_n^{-1}(t) \leq x$ provided $x > z$ is a continuity point of F.
Thus $\overline{\lim}\, F_n^{-1}(t) \leq z$ since there are continuity points x that $\downarrow z$.

Thus $F_n^{-1}(t) \to F^{-1}(t)$ for all but a countable number of t's. Thus the definition (a) of X_n^* works. □

Proposition 3.6. (Continuous mapping or Mann - Wald theorem; restatement and proof of proposition 2.8). Suppose that $g : R \to R$ is continuous a.s. P_X. Then:

A. If $X_n \to_{a.s.} X$, then $g(X_n) \to_{a.s.} g(X)$.

B. If $X_n \to_p X$, then $g(X_n) \to_p g(X)$.

C. If $X_n \to_d X$, then $g(X_n) \to_d g(X)$.

Proof. A. Let N_1 be the null set such that $X_n(\omega) \to X_0(\omega)$ for all $\omega \notin N_1$; and let N_2 be the null set such g is continuous at $X_0(\omega)$ for all $\omega \notin N_2$. Then for $\omega \notin N_1 \cup N_2$ we have $g(X_n(\omega)) \to g(X_0(\omega))$. But $P(N_1 \cup N_2) \leq P(N_1) + P(N_2) = 0$, and the convergence asserted in A holds.

C. Replace X_n, X_0 by X_n^*, X_0^* of the Skorokhod proposition 3.5. Thus

(a) $$g(X_n) \stackrel{d}{=} g(X_n^*) \to_{a.s.} g(X_0^*) \stackrel{d}{=} g(X_0).$$

Since $\to_{a.s.}$ implies $\to_p$, which in turn implies $\to_d$, (a) implies that $g(X_n) \to_d g(X_0)$.

B. Probably the easist way to do this one is via theorem 2.1.6, part G. Then the argument is the same as that in A, but for a subsequence. □

Remark 1. Proposition 3.5 remains true for random vectors in R^k and, still more generally, for convergence in law (weak convergence) of random elements in a separable metric space. See Billingsley (1986), *Probability and Measure*, page 399, for the first; Billingsley (1971), *Weak Convergence of Measures: Applications in Probability*, theorem 3.3, page 7, for the second. The original paper is Skorokhod (1956) where the separable metric space case was treated immediately.

Proposition 3.7. (Helly - Bray theorem). If $X_n \to_d X_0$ and g is bounded and continuous (a.s. P_X), then $Eg(X_n) \to Eg(X_0)$.

Proof. For the rv's X_n^* of proposition 3.5, it follows from A of proposition 3.6 that $g(X_n^*) \to_{a.s.} g(X_0^*)$. Thus by equality in laws guaranteed by the construction of proposition 3.5 and the dominated convergence theorem,

$$Eg(X_n) = Eg(X_n^*) \to Eg(X_0^*) = Eg(X_0) .$$

□

Remark 2. If $Eg(X_n) \to Eg(X)$ for all bounded continous functions g , then $X_n \to_d X_0$. (Proof: box in the indicator function $1_{(-\infty, t]}$ by the bounded continuous functions g_+ , g_- defined by connecting $(t, 1)$ to $(t+\epsilon, 0)$ linearly and $(t-\epsilon, 1)$ to $(t, 0)$ linearly, respectively.) This gives a way of defining $\to_d$ more generally:

Definition 3.8. Suppose that $X_n , n \geq 0$ take values in the complete separable metric space (M, d) . Then we say that X_n *converges in law or distribution to* X_0 , and we write $X_n \to_d X_0$ or $X_n \Longrightarrow X_0$, if

$$Eg(X_n) \to Eg(X_0) \qquad \text{for all } g \in C_b(M) ;$$

here $C_b(M)$ denotes the collection of all bounded continuous functions from M to R .

Proposition 3.9. If $X_n \to_d X_0$, then $E|X_0| \leq \liminf_{n\to\infty} E|X_n|$.

Proof. For X_n^* of proposition 3.5, $X_n^* \to_{a.s.} X_0^*$, so, by the equality in law of proposition 3.5 and Fatou's lemma

$$E|X_0| = E|X_0^*| = E(\liminf_n |X_n^*|) \leq \liminf_n E|X_n^*| = \liminf_n E|X_n| .$$

□

Proposition 3.10. (Slutsky's theorem; restatement and proof of proposition 2.2.9). If $A_n \to_p a$, $B_n \to_p b$, and $Z_n \to_d Z$, then $A_n Z_n + B_n \to_d aZ + b$.

Proof. Now $A_n \to_p a$, $B_n \to_p b$, and $Z_n \to_d Z$ where a, b are constants, implies that $(Z_n, A_n, B_n) \to_d (Z, a, b)$ in R^3 . Hence by the R^3 version of Skorokhod's theorem, there exists a sequence $(Z_n^*, A_n^*, B_n^*) \stackrel{d}{=} (Z_n, A_n, B_n)$ such that $(Z_n^*, A_n^*, B_n^*) \to_{a.s.} (Z^*, a, b)$. Hence

(a) $$A_n Z_n + B_n \stackrel{d}{=} A_n^* Z_n^* + B_n^* \rightarrow_{a.s.} aZ^* + b =^d aZ + b .$$

Since $\rightarrow_{a.s.}$ implies $\rightarrow_p$, which in turn implies $\rightarrow_d$, (a) yields the desired conclusion. □

Proposition 3.11. (The $g'-$ theorem, or the delta - method). If $X_n \equiv a_n(Z_n - b) \rightarrow_d X$ where $a_n \rightarrow \infty$, and g is differentiable at b , then

$$a_n\{g(Z_n) - g(b)\} \rightarrow_d g'(b)X .$$

Proof. By proposition 3.5, there exists a sequence $X_n^* \stackrel{d}{=} X_n$ such that $X_n^* \rightarrow_{a.s.} X$. Setting $Z_n^* \equiv (X_n^*/a_n) + b$, it follows that $Z_n^* \stackrel{d}{=} Z_n$ and $X_n^* = a_n(Z_n^* - b) \rightarrow_{a.s.} X^*$. Since $a_n \rightarrow \infty$ we have $Z_n^* \rightarrow_{a.s.} b$. Thus

$$\begin{aligned} a_n(g(Z_n) - g(b)) &\stackrel{d}{=} a_n(g(Z_n^*) - g(b)) \\ &= \frac{g(Z_n^*) - g(b)}{Z_n^* - b} a_n(Z_n^* - b) \\ &\rightarrow_{a.s.} g'(b)X^* \stackrel{d}{=} g'(b)X . \end{aligned}$$

Since $\rightarrow_{a.s.}$ implies $\rightarrow_p$, which in turn implies $\rightarrow_d$, the conclusion follows. □

4. Empirical Distribution Functions, Empirical Measures, and Empirical Processes

Suppose that $\xi_1, \cdots, \xi_n, \cdots$ are iid Uniform$(0,1)$. Their *empirical df* is

$$\mathbb{G}_n(t) \equiv \frac{1}{n}\sum_{i=1}^{n} 1_{[0,t]}(\xi_i) \qquad \text{for} \quad 0 \le t \le 1 \tag{1}$$

$$= \frac{k}{n} \qquad \text{for} \quad \xi_{n:k} \le t < \xi_{n:k+1}, \quad 0 \le k \le n \tag{2}$$

where $0 \equiv \xi_{n:0} \le \xi_{n:1} \le \cdots \le \xi_{n:n} \le \xi_{n:n+1} \equiv 1$ are the order statistics. The *uniform empirical process* is defined by

$$\mathbb{U}_n(t) \equiv \sqrt{n}\,[\mathbb{G}_n(t) - t] \qquad \text{for} \quad 0 \le t \le 1\,. \tag{3}$$

The inverse function $\mathbb{G}_n^{-1}$ of $\mathbb{G}_n$ is the *uniform quantile function;* Thus $\mathbb{G}_n^{-1}(t) = \xi_{n:i}$ for $(i-1)/n < t \le i/n$, $i = 1, \cdots, n$. The *uniform quantile process* $\mathbb{V}_n$ is defined by

$$\mathbb{V}_n(t) \equiv \sqrt{n}\,[\mathbb{G}_n^{-1}(t) - t] \qquad \text{for} \quad 0 \le t \le 1\,. \tag{4}$$

Note that

$$\mathbb{G}_n(t) \sim n^{-1}\,\mathit{Binomial}(n, t) \sim (t, t(1-t)/n) \qquad \text{for} \quad 0 \le t \le 1\,, \tag{5}$$

so that

(6) $\qquad \mathbb{U}_n(t) \;\sim\; (\,0\,, t(1-t)\,) \qquad \text{for} \quad 0 \,\le\, t \,\le\, 1\,.$

In fact

(7) $\qquad Cov[\,1_{[0,s]}(\xi_i)\,, 1_{[0,t]}(\xi_i)\,] \;=\; s \wedge t \;-\; st \qquad \text{for} \;\; 0 \,\le\, s, t \,\le\, 1\,.$

Moreover, applying the multivariate CLT to $(\,1_{[0,s]}(\xi_i)\,, 1_{[0,t]}(\xi_i)\,)$, it is clear that

(8) $$(\,\mathbb{U}_n(s)\,, \mathbb{U}_n(t)\,) \;\to_d\; N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} s(1-s) & s \wedge t - st \\ s \wedge t - st & t(1-t) \end{pmatrix}) \qquad \text{as} \quad n \to \infty\,,$$

for $0 \,\le\, s, t \,\le\, 1$.

We define $\{\,\mathbb{U}(t) :\; 0 \,\le\, t \,\le\, 1\,\}$ to be a *Brownian bridge* if it is a Gaussian process on $(R_{[0,1]}\,, \mathbf{B}_{[0,1]}\,)$ having

(9) $\qquad E\,\mathbb{U}(t) \;=\; 0 \qquad \text{and} \qquad Cov[\mathbb{U}(s)\,, \mathbb{U}(t)] \;=\; s \wedge t \;-\; st$

for all $0 \,\le\, s, t \,\le\, 1$. The process $\mathbb{U}$ exists, and has sample functions $\mathbb{U}(\cdot\,, \omega)$ which are continuous for a.e. ω .

Of course the bivariate result (8) immediately extends to the multivariate $(\mathbb{U}_n(t_1)\,, \cdots, \mathbb{U}_n(t_k))$: for any $1 \le k < \infty$ and any $t_1\,, \cdots, t_k \,\in\, [0,1]$,

$$(\mathbb{U}_n(t_1)\,, \cdots, \mathbb{U}_n(t_k)) \;\to_d\; (\mathbb{U}(t_1)\,, \cdots, \mathbb{U}(t_k)) \;\sim\; N_k(0, (t_i \wedge t_j - t_i t_j))\,.$$

Thus we have convergence of all the *finite - dimensional distributions* of $\mathbb{U}_n$ to those of $\mathbb{U}$, and we write:

(10) $\qquad \mathbb{U}_n \;\to_{f.d.}\; \mathbb{U} \qquad \text{as} \qquad n \,\to\, \infty\,.$

The inference we would like to draw from (10) is that $g(\mathbb{U}_n) \;\to_d\; g(\mathbb{U})$ as $n \,\to\, \infty$ as $n \,\to\, \infty$ for various functionals such as $g \,\equiv\, ||\cdot||$. The conclusion (10) is not strong enough to imply this, but (10) can be strengthened to a result that does. The Mann - Wald theorem suggests $g(\mathbb{U}_n) \to_d g(\mathbb{U})$ should be true for "continuous" functionals g , and this raises the question of what metric should be used to define "continuous".

The Empirical Process on $\mathbb{R}$

Let $X_1, \cdots, X_n$ be iid F with order statistics $X_{n:1} \le \cdots \le X_{n:n}$. Their *empirical df* $\mathbb{F}_n$ is defined by

(11) $$\mathbb{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty, x]}(X_i) \qquad \text{for } -\infty < x < \infty .$$

The *empirical process* is defined to be $\sqrt{n}\,(\mathbb{F}_n - F)$.

It is extremely convenient to suppose that the X_i^*'s are defined by

(12) $$X_i^* = F^{-1}(\xi_i) \qquad 1 \le i \le n, \qquad \text{for the } \xi_i's \text{ of (1);}$$

proposition 2.3.4 shows that these X_i^*'s are indeed iid F. Recall from proposition 2.3.4 that also

(13) $$1_{[X_i^* \le x]} = 1_{[\xi_i \le F(x)]} \qquad \text{on} \quad (-\infty, \infty) \qquad \text{a.s.}$$

for these particular X_i^*'s. Thus for these X_i^*'s we have

(14) $$\mathbb{F}_n^* = \mathbb{G}_n(F) \qquad \text{on} \quad (-\infty, \infty) \qquad \text{a.s.}$$

and

(15) $$\sqrt{n}\,[\mathbb{F}_n^* - F] = \mathbb{U}_n(F) \qquad \text{on} \quad (-\infty, \infty) \qquad a.s..$$

Note from (10) and (15) that

(16) $$\sqrt{n}\,(\mathbb{F}_n - F) \to_{f.d.} \mathbb{U}(F) \qquad \text{as} \quad n \to \infty .$$

Theorem 4.1. (Glivenko - Cantelli).

(17) $$||\mathbb{G}_n - G||_\infty \equiv \sup_{0 \le t \le 1} |\mathbb{G}_n(t) - t| \to_{a.s.} 0$$

and

(18) $$||\mathbb{F}_n - F||_\infty \equiv \sup_{-\infty < x < \infty} |\mathbb{F}_n(x) - F(x)| \to_{a.s.} 0$$

as $n \to \infty$.

Proof. Since

(a) $$||\mathbb{F}_n - F||_\infty = ||\mathbb{G}_n(F) - F||_\infty \le ||\mathbb{G}_n - G||_\infty ,$$

with "=" if F is continuous, it suffices to prove the first part.

Fix a large integer M. Then

$$\begin{aligned} ||\mathbb{G}_n - G||_\infty &= \max_{1 \le i \le M} \sup_{(i-1)/M \le t \le i/M} |\mathbb{G}_n(t) - t| \\ &= \max_{1 \le i \le M} \Big\{ \sup_{(i-1)/M \le t \le i/M} (\mathbb{G}_n(t) - t) \end{aligned}$$

$$\bigvee \sup_{(i-1)/M \,\le\, t \,\le\, i/M} (t \;-\; \mathbb{G}_n(t))\}$$

$$\le \max_{1 \,\le\, i \,\le\, M} \{(\mathbb{G}_n(\frac{i}{M}) \;-\; \frac{i-1}{M}) \;\bigvee\; (\frac{i}{M} \;-\; \mathbb{G}_n(\frac{i-1}{M}))\}$$

$$\le \max_{1 \,\le\, i \,\le\, M} \{(\mathbb{G}_n(\frac{i}{M}) \;-\; \frac{i}{M}) \;\bigvee\; (\frac{i-1}{M} \;-\; \mathbb{G}_n(\frac{i-1}{M}))\} \;+\; \frac{1}{M}$$

$$\to_{a.s.} \; 0 \;+\; \frac{1}{M}$$

since $\mathbb{G}_n(\frac{i}{M}) \to_{a.s.} \frac{i}{M} = G(\frac{i}{M})$, $1 \le i \le M$. But M was arbitrary; hence $||\mathbb{G}_n - G||_\infty \to_{a.s.} 0$. $\square$

Theorem 4.2. (Special empirical processes). There exists a (sequence of) Brownian bridge processes $\mathbb{U} = \mathbb{U}^{(n)}$ and a triangular array of row independent Uniform$(0,1)$ rv's $\xi_{n1}, \cdots, \xi_{nn}$, $n \ge 1$ all defined on a common probability space $(\Omega, \mathbf{A}, P)$, such that if $\mathbb{U}_n$ denotes the empirical process of $\xi_{n1}, \cdots, \xi_{nn}$, then

(19) $$||\mathbb{U}_n - \mathbb{U}||_\infty \equiv \sup_{0 \,\le\, t \,\le\, 1} |\mathbb{U}_n(t) - \mathbb{U}(t)| \to_{a.s.} 0 .$$

Thus

(20) $$||\sqrt{n}(\mathbb{F}_n^* - F) - \mathbb{U}(F)||_\infty \to_{a.s.} 0 \qquad \text{as} \quad n \to \infty .$$

Corollary 4.3. (Donsker's theorem). If $g : D[0,1] \to R$ is $||\;||_\infty$ − continuous, then $g(\mathbb{U}_n) \to_d g(\mathbb{U})$.

Example 1. (Kolmogorov's statistic). If F is continuous,

$$||\sqrt{n}(\mathbb{F}_n - F)||_\infty \stackrel{d}{=} ||\mathbb{U}_n(F)||_\infty = ||\mathbb{U}_n||_\infty \to_d ||\mathbb{U}||_\infty .$$

It is known, via reflection methods (see Shorack and Wellner (1986), pages 33 - 42, 142), that

(21) $$P(\, ||\mathbb{U}|| > \lambda) = 2 \sum_{k=1}^{\infty} (-1)^{k+1} \exp(-2k^2\lambda^2) \qquad \text{for} \quad \lambda > 0 .$$

Example 2. (Kolmogorov's one-sided statistic). If F is continuous,

$$||\sqrt{n}(\mathbb{F}_n - F)^+||_\infty \stackrel{d}{=} ||\mathbb{U}_n^+(F)||_\infty = ||\mathbb{U}_n^+||_\infty \to_d ||\mathbb{U}^+||_\infty .$$

It is known, (see Shorack and Wellner, (1986), pages 37 and 142), that

(22) $$P(\, ||\mathbb{U}^+|| > \lambda) = \exp(-2\lambda^2) \qquad \text{for} \quad \lambda > 0 .$$

Example 3. (Birnbaum's statistic). If F is continuous,

$$\int_{-\infty}^{\infty} \sqrt{n}(I\!F_n - F)\,dF \stackrel{d}{=} \int_{-\infty}^{\infty} I\!U_n(F)\,dF = \int_0^1 I\!U_n(t)\,dt \to_d \int_0^1 I\!U(t)\,dt$$

Now $\int_0^1 I\!U(t)dt$ is a linear combination of normal rv's, and hence is normal. It has expected value 0 since $E\,I\!U(t) = 0$ for each fixed t. Furthermore

$$\begin{aligned} E(\int_0^1 I\!U(t)dt)^2 &= E(\int_0^1 I\!U(s)ds)(\int_0^1 I\!U(t)dt) \\ &= \int_0^1\int_0^1 EI\!U(s)I\!U(t)\,dsdt \\ &= \int_0^1\int_0^1 (s \wedge t - st)\,ds\,dt = \frac{1}{12}\,. \end{aligned}$$

Thus $\int_0^1 I\!U(t)dt \sim N(0, \frac{1}{12})$.

Example 4. (Cramér - von Mises statistic). If F is continuous,

$$\begin{aligned} \int_{-\infty}^{\infty} \{\sqrt{n}(I\!F_n - F)\}^2\,dF &\stackrel{d}{=} \int_{-\infty}^{\infty} \{I\!U_n(F)\}^2\,dF \\ &= \int_0^1 \{I\!U_n(t)\}^2\,dt \to_d \int_0^1 \{I\!U(t)\}^2\,dt\,. \end{aligned}$$

It is known that

$$\int_0^1 \{I\!U(t)\}^2\,dt \stackrel{d}{=} \sum_{j=1}^{\infty} \frac{1}{j^2\pi^2} Z_j^2$$

where the Z_j 's are iid $N(0,1)$; this distribution has been tabled. See Shorack and Wellner (1986), page 147.

Example 5. (Anderson - Darling statistic). If F is continuous,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\{\sqrt{n}(I\!F_n - F)\}^2}{F(1-F)}\,dF &\stackrel{d}{=} \int_{-\infty}^{\infty} \frac{\{I\!U_n(F)\}^2}{F(1-F)}\,dF \\ &= \int_0^1 \frac{\{I\!U_n(t)\}^2}{t(1-t)}\,dt \to_d \int_0^1 \frac{\{I\!U(t)\}^2}{t(1-t)}\,dt\,. \end{aligned}$$

It is known that

$$\int_0^1 \frac{\{I\!U(t)\}^2}{t(1-t)}\,dt \stackrel{d}{=} \sum_{j=1}^{\infty} \frac{1}{j(j+1)} Z_j^2$$

where the Z_j 's are iid $N(0,1)$, and this distribution has been tabled; see Shorack and Wellner (1986), page 148.

General Empirical Measures and Processes

Now suppose that $X_1, X_2, \cdots$ are iid P on the measurable space $(A, \mathbf{A})$. We let $I\!P_n$ denote the *empirical measure* of the first n X_i's:

$$I\!P_n \equiv \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i} ;$$

thus for a set $B \in \mathbf{A}$,

$$I\!P_n(B) \equiv \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \frac{1}{n} \{\text{Number of } X_i \text{ in } B,\ i = 1, \cdots, n\} .$$

Note that when $A = R$, so that the X_i's are real - valued, and $B = (-\infty, x]$ for $x \in R$, then

$$I\!P_n(B) = I\!P_n((-\infty, x]) = I\!F_n(x) ,$$

the empirical distribution function of the X_i's at x.

From the strong law of large numbers and the ordinary CLT it follows that for any fixed set $B \in \mathbf{A}$

(23) $$I\!P_n(B) \rightarrow_{a.s.} P(B)$$

and

(24) $$X\!\!X_n(B) \equiv \sqrt{n}(I\!P_n(B) - P(B)) \rightarrow_d X\!\!X(B) \sim N(0, P(B)(1 - P(B))) .$$

Similarly, for any finite number of sets $B_1, \cdots, B_k \in \mathbf{A}$ we have, by the SLLN and the multivariate CLT,

(25) $$I\!P_n(\underline{B}) \equiv (I\!P_n(B_1), \cdots, I\!P_n(B_k))$$

(26) $$\rightarrow_{a.s.} (P(B_1), \cdots, P(B_k)) \equiv P(\underline{B})$$

and

$$Z\!\!Z_n(\underline{B}) \equiv \sqrt{n}(I\!P_n(\underline{B}) - P(\underline{B})) \rightarrow_d Z\!\!Z(\underline{B}) \sim N_k(0, C)$$

where C has i, j th entry $C_{ij} = P(B_i \cap B_j) - P(B_i)P(B_j)$. If $Z\!\!Z$ is a mean zero Gaussian process indexed by some class of subsets $\mathbf{C}$ of A with

$$Cov[Z\!\!Z(B), Z\!\!Z(C)] = P(B \cap C) - P(B)P(C), \quad \text{for } B, C \in \mathbf{C},$$

and with sample paths which are a.s. continuous with respect to the pseudo - metric $\rho = \rho_P$ given by $\rho^2(B, C) = Var(1_B(X) - 1_C(X))$, then we say that $Z\!\!Z$ is a P *− Brownian bridge process indexed by* $\mathbf{C}$.

The question of interest is: for what classes $\mathbf{C}$ of subsets of A, $\mathbf{C} \subset \mathbf{A}$, can we make the convergences in (23) and (24) uniform over $\mathbf{C}$? These are the kinds of questions that modern empirical process theory can answer; see e.g. Pollard (1984) for some of this. Here is one partial answer to the question:

Theorem 4.3. (Vapnik and Chervonenkis (1971); Dudley (1978)). Suppose that $\mathbf{C}$ is a *Vapnik - Chervonenkis class* of subsets of $\mathbf{A}$ (satisfying a little bit of measurability hypotheses). Then

$$(27) \qquad ||\mathbb{P}_n - P||_{\mathbf{C}} \equiv \sup_{C \in \mathbf{C}} |\mathbb{P}_n(C) - P(C)| \rightarrow_{a.s.} 0$$

as $n \rightarrow \infty$. Furthermore we can construct the X_i's and (a sequence of) $P-$ Brownian bridge process(es) $\mathbb{Z} = \mathbb{Z}^{(n)}$ on a common probability space so that

$$(28) \qquad ||\mathbb{Z}_n - \mathbb{Z}||_{\mathbf{C}} \rightarrow_p 0$$

as $n \rightarrow \infty$.

5. The Partial Sum Process and Brownian Motion

We define $\{ \mathbb{B}(t) : 0 \le t \le 1 \}$ to be *Brownian motion* if $\mathbb{B}$ is a Gaussian process on $R_{[0,1]}, \mathbf{B}_{[0,1]})$ having

$$(1) \qquad E\,\mathbb{B}(t) = 0 \qquad \text{and} \qquad Cov[\mathbb{B}(s), \mathbb{B}(t)] = s \wedge t$$

for all $0 \le s, t \le 1$. Theorem 1.3.3(i) shows that these distributions are "consistent"; thus Kolmogorov's theorem (see e.g. Billingsley's (1986) *Probability and Measure*, page 506ff) shows that the process $\mathbb{B}$ exists. Note that (1) and normality imply that

$$(2) \qquad \mathbb{B} \quad \text{has stationary independent increments} .$$

Exercise 5.1. Suppose that $\mathbb{U}$ is a Brownian bridge and $Z \sim N(0,1)$ is independent of $\mathbb{U}$. Let

$$(3) \qquad \mathbb{B}(t) \equiv \mathbb{U}(t) + t\,Z \qquad \text{for} \quad 0 \le t \le 1 .$$

Show that $\mathbb{B}$ is a Brownian motion.

Exercise 5.2. Suppose $\mathbb{B}$ is a Brownian motion. Show that

$$(4) \qquad \mathbb{U}(t) \equiv \mathbb{B}(t) - t\,\mathbb{B}(1) \qquad \text{for} \quad 0 \le t \le 1$$

is a Brownian bridge.

Let $X_1, \cdots, X_n$ be iid $(0,1)$ rv's, and let $X_0 \equiv 0$. We define the *partial sum process* $\mathbb{B}_n$ by

$$(5) \qquad \mathbb{B}_n(t) \equiv \mathbb{B}_n(\frac{k}{n}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{k} X_i \qquad \text{for} \quad \frac{k}{n} \le t < \frac{k+1}{n} ,$$

for $0 \le k < \infty$.

Note that

$$(6) \qquad Cov[\mathbb{B}_n(j/n)\,,\,\mathbb{B}_n(k/n)\,] \;=\; \frac{1}{n}\sum_{i=1}^{j}\sum_{i'=1}^{k} Cov[\,X_i\,,X_{i'}\,]$$

$$= \frac{1}{n}\sum_{i=1}^{j\wedge k} Var[\,X_i\,] \;=\; \frac{j\wedge k}{n}$$

$$(7) \qquad \to\; s\wedge t \qquad \text{if } \; j/n \to s \quad \text{and} \quad k/n \to t\,.$$

Also, for $0 \le t \le 1$,

$$(8) \qquad \mathbb{B}_n(t) \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} X_i \;=\; \sqrt{\frac{[nt]}{n}}\,\frac{1}{\sqrt{[nt]}}\sum_{i=1}^{[nt]} X_i \;\to_d\; \sqrt{t}\,N(0,1) \;\sim\; N(0,t)$$

for $0 \le t \le 1$ by the CLT and Slutsky's theorem 2.2.9 (or 2.3.10). This suggests that

$$(9) \qquad \mathbb{B}_n \;\to_{f.d.}\; \mathbb{B} \qquad \text{as } \; n\to\infty\,.$$

This will be verified in exercise 5.3. Much more is true: $\mathbb{B}_n \Longrightarrow \mathbb{B}$ as processes in $D[0,1]$, and hence $g(\mathbb{B}_n) \to_d g(\mathbb{B})$ for continuous functionals g .

Exercise 5.3. Show that (9) holds: i.e. for any fixed $t_1,\cdots,t_k \in [0,1]^k$, $(\mathbb{B}_n(t_1)\,,\cdots,\mathbb{B}_n(t_k)) \;\to_d\; N_k(0\,,(t_i\wedge t_j))$.

Existence of Brownian Motion and Brownian Bridge as Continuous Processes on $C[0,1]$

The aim of this subsection is to convince you that both Brownian motion and Brownian bridge exist as *continuous Gaussian processes* on $[0,1]$, and that we can then extend the definition of Brownian motion to $[0\,,\infty)$.

Definition 5.1. *Brownian motion* (or *standard Brownian motion*, or, a *Wiener process*) $\mathbb{B}$ is a Gaussian process with continuous sample functions and:

(i) $\mathbb{B}(0) = 0$.

(ii) $E\mathbb{B}(t) = 0\,,\; 0 \le t \le 1$.

(iii) $E\{\mathbb{B}(s)\mathbb{B}(t)\} \;=\; s\wedge t\,,\quad 0 \le s\,,t \le 1$.

Definition 5.2. A *Brownian bridge process* $\mathbb{U}$ is a Gaussian process with continuous sample functions and:

(i) $E\mathbb{U}(t) = 0\,,\; 0 \le t \le 1$.

(ii) $E\{\mathbb{U}(s)\mathbb{U}(t)\} \;=\; s\wedge t \,-\, s\,t,\quad 0 \le s\,,t \le 1$.

Theorem 5.1. Brownian motion $\mathbb{B}$ and Brownian bridge $\mathbb{U}$ exist.

Proof. We first construct a Brownian bridge process $I\!U$. Let

$$h_{00}(t) \equiv h(t) \equiv \begin{cases} t & 0 \le t \le 1/2 \\ 1-t & 1/2 \le t \le 1 \\ 0 & elsewhere \end{cases} .$$

Let

$$h_{nj}(t) \equiv 2^{-n/2}\, h(2^n t - j)\,, \quad j = 0, \cdots, 2^n - 1\,.$$

For example, $h_{10}(t) = 2^{-1/2}h(2t)$, $h_{11}(t) = 2^{-1/2}h(2t-1)$, while,

$$h_{20}(t) = 2^{-1}h(4t)\,, \quad h_{21}(t) = 2^{-1}h(4t-1)\,,$$

$$h_{22}(t) = 2^{-1}h(4t-2)\,, \quad h_{23}(t) = 2^{-1}h(4t-3)\,.$$

Note that $|h_{nj}(t)| \le 2^{-n/2}\, 2^{-1}$.

The functions $\{h_{nj}\}_{j=0\,,\,n\ge 0}^{2^n-1}$ are called the *Schauder functions*; they are integrals of the orthonormal (with respect to Lebesgue measure on $[0,1]$) family of functions $\{g_{nj}\}_{j=0\,,\,n\ge 0}^{2^n-1}$ called the *Haar functions* defined by

$$g_{00}(t) \equiv g(t) \equiv 2 1_{[0,1/2]}(t) - 1\,,$$

$$g_{nj}(t) \equiv 2^{n/2}\, g_{00}(2^n t - j)\,, \quad j = 0, \cdots, 2^n - 1, \quad n \ge 1\,.$$

Thus

$$\int_0^1 g_{nj}^2\, d\lambda = 1\,, \qquad \int_0^1 g_{nj}\, g_{n'j'}\, d\lambda = 0\,, \qquad \text{if} \quad n \ne n'\,, \quad \text{or} \quad j' \ne j\,,$$

and

$$h_{nj}(t) = \int_0^t g_{nj}(u) du = \int_0^t g_{nj}\, d\lambda\,, \quad 0 \le t \le 1\,.$$

Furthermore, the family $\{g_{nj}\}_{j=0\,,\,n\ge 0}^{2^n-1} \cup \{g(\,\cdot\,/2)\}$ is complete; any $f \in L^2[0,1]$ has an expansion in terms of the g 's.

Now let $\{X_{nj}\}_{j=0\,,\,n\ge 0}^{2^n-1}$ be independent identically distributed $N(0,1)$ random variables; if we wanted, we could construct all these random variables on the probability space $([0,1]\,, \mathbf{B}([0,1])\,, \lambda)$. Define

$$V_n(t\,,\omega) = \sum_{j=0}^{2^n-1} h_{nj}(t)\, X_{nj}(\omega)\,,$$

$$U_m(t\,,\omega) = \sum_{n=0}^{m} V_n(t\,,\omega)\,.$$

For $m > k$,

$$|U_m(t,\omega) - U_k(t,\omega)| = |\sum_{n=k+1}^{m} V_n(t,\omega)| \le \sum_{n=k+1}^{m} |V_n(t,\omega)|$$

where

$$|V_n(t,\omega)| \;\le\; \sum_{j=0}^{2^n-1} |h_{nj}(t)||X_{nj}(\omega)| \;\le\; 2^{-n/2+1} \max_{0\le j\le 2^n-1} |X_{nj}(\omega)|$$

since the h_{nj}, $j = 0, \cdots, 2^n - 1$ are $\ne 0$ on disjoint t intervals.

Now $P(X_{nj} \ge z) = 1 - \Phi(z) \le z^{-1}\phi(z)$ for $z > 0$ (Mill's ratio), so that

$$P(|X_{nj}| \ge 2\sqrt{n}) = 2\,P(X_{nj} \ge 2\sqrt{n}) \le \frac{2}{\sqrt{2\pi}}(2\sqrt{n})^{-1}e^{-2n}\,.$$

Hence

$$P(\max_{0\le j\le 2^n-1}|X_{nj}| \ge 2\sqrt{n}) \le 2^n P(|X_{00}| \ge 2\sqrt{n}) \le \frac{2^n}{\sqrt{2\pi}} n^{-1/2} e^{-2n}\,;$$

since this is a term of a convergent series, by the Borel-Cantelli lemma, $\max_{0\le j\le 2^n-1}|X_{nj}| \ge 2\sqrt{n}$ occurs finitely often with probability one: i.e. except on a null set, for all ω there is an $N \equiv N(\omega)$ such that $\max_{0\le j\le 2^n-1}|X_{nj}(\omega)| < 2\sqrt{n}$ for all $n > N(\omega)$. Hence

$$\sup_{0\le t\le 1}|U_m(t) - U_k(t)| \le \sum_{n=k+1}^{m} 2^{-n/2} n^{1/2} \quad \downarrow\; 0$$

for all $k, m \ge N' \ge N(\omega)$. Thus $U_m(t,\omega)$ converges uniformly as $m \to \infty$ with probability one to (the necessarily continuous) function

$$\mathbb{U}(t,\omega) \equiv \sum_{n=0}^{\infty} V_n(t,\omega)\,.$$

Define $\mathbb{U} \equiv 0$ on the exceptional set. Then $\mathbb{U}$ is continuous for all ω.

Now $\{\mathbb{U}(t) : 0 \le t \le 1\}$ is clearly a Gaussian process since it is the sum of Gaussian processes. We now show that $\mathbb{U}$ is in fact a Brownian bridge: by formal calculation (it remains only to justify the interchange of summation and expectation),

$$\begin{aligned}
E\{\mathbb{U}(s)\mathbb{U}(t)\} &= E\{\sum_{n=0}^{\infty} V_n(s) \sum_{m=0}^{\infty} V_m(t)\} \\
&= \sum_{n=0}^{\infty} E\{V_n(s)\,V_n(t)\} \\
&= E\{\sum_{j=0}^{2^n-1} X_{nj}\int_0^s g_{nj} \sum_{k=0}^{2^n-1} X_{nk}\int_0^t g_{nk}\} \\
&= \sum_{n=0}^{\infty}\sum_{j=0}^{2^n-1}\int_0^s g_{nj}\int_0^t g_{nj}
\end{aligned}$$

$$= \sum_{n=0}^{\infty} \sum_{j=0}^{2^n-1} \int_0^1 1_{[0,s]} g_{nj} \int_0^1 1_{[0,t]} g_{nj} \ + st \ - st$$

$$= \int_0^1 1_{[0,s]}(u) 1_{[0,t]}(u) \, du \ - st$$

$$= s \bigwedge t \ - st$$

where the next to last equality follows from Parseval's identity. Thus $\mathbb{U}$ is Brownian bridge.

Now let X be one additional $N(0,1)$ rv independent of all the others used in the construction, and define

$$\mathbb{B}(t) \equiv \mathbb{U}(t) + tX = \sum_{n=0}^{\infty} V_n(t) + tX .$$

Then $\mathbb{B}$ is also Gaussian with 0 mean and

$$\begin{aligned} Cov(\mathbb{B}(s), \mathbb{B}(t)) &= Cov(\mathbb{U}(s) + sX, \mathbb{U}(t) + tX) \\ &= Cov(\mathbb{U}(s), \mathbb{U}(t)) + s\,t\,Var(X) \\ &= s \bigwedge t - st + st = s \bigwedge t . \end{aligned}$$

Thus $\mathbb{B}$ is Brownian motion; since U has continuous sample paths, so does $\mathbb{B}$. □

Exercise 5.4. Graph the first few g_{nj}'s and h_{nj}'s.

Exercise 5.5 Justify the interchange of expectation and summation used in the proof. [Hint: use the Tonelli part of Fubini's theorem.]

Exercise 5.6. Let $\mathbb{U}$ be a Brownian bridge process. For $0 \le t < \infty$ define a process $\mathbb{B}$ by

$$\mathbb{B}(t) \equiv (1+t)\,\mathbb{U}\left(\frac{t}{1+t}\right) .$$

Show that $\mathbb{B}$ is a Brownian motion process on $[0, \infty)$.

6. Quantiles and Quantile Processes

Let $X_1, \cdots, X_n$ be iid real-valued rv's with distribution function F, and let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ denote the *order statistics.* For $t \in (0,1)$, let

$$\mathbb{F}_n^{-1}(t) \equiv \inf\{x : \mathbb{F}_n(x) \ge t\}$$

so that

$$\mathbb{F}_n^{-1}(t) = X_{(i)} \qquad \text{for } \frac{i-1}{n} < t \le \frac{i}{n},\ i = 1, \cdots, n .$$

Let $\xi_1, \cdots, \xi_n$ be iid Uniform$(0,1)$ rv's and let $0 \le \xi_{(1)} \le \cdots \le \xi_{(n)} \le 1$ denote their order statistics. Thus, with $\mathbb{G}_n^{-1}(t) \equiv \inf\{x : \mathbb{G}_n(x) \ge t\}$,

$$\mathbb{G}_n^{-1}(t) = \xi_{(i)} \qquad \text{for } \frac{i-1}{n} < t \le \frac{i}{n},\ i = 1, \cdots, n .$$

Now

$$(X_1^*, \cdots, X_n^*) \equiv (F^{-1}(\xi_1), \cdots, F^{-1}(\xi_n)) \stackrel{d}{=} (X_1, \cdots, X_n),$$

so

$$(X_{(1)}^*, \cdots, X_{(n)}^*) \equiv (F^{-1}(\xi_{(1)}), \cdots, F^{-1}(\xi_{(n)})) \stackrel{d}{=} (X_{(1)}, \cdots, X_{(n)}).$$

Hence

$$\mathbb{F}_n^{-1}(\cdot) \stackrel{d}{=} F^{-1}(\mathbb{G}_n^{-1}(\cdot)),$$

and to study $\mathbb{F}_n^{-1}$ it suffices to study $\mathbb{G}_n^{-1}$.

Proposition 6.1.

$$||\mathbb{G}_n^{-1} - I|| \equiv \sup_{0 \le t \le 1} |\mathbb{G}_n^{-1}(t) - t| = ||\mathbb{G}_n - I|| \to_{a.s.} 0,$$

and hence, if F^{-1} is continuous on $[a,b] \subset [0,1]$, then

$$||\mathbb{F}_n^{-1} - F^{-1}||_a^b \equiv \sup_{a \le t \le b} |\mathbb{F}_n^{-1}(t) - F^{-1}(t)| \to_{a.s.} 0 .$$

Proof. $||\mathbb{G}_n^{-1} - I|| = ||\mathbb{G}_n - I||$ by inspection of the graphs. Thus

$$||\mathbb{F}_n^{-1} - F^{-1}||_a^b \stackrel{d}{=} ||F^{-1}(\mathbb{G}_n^{-1}) - F^{-1}(I)||_a^b \to_{a.s.} 0$$

since F^{-1} is uniformly continuous on $[a,b]$ and $||\mathbb{G}_n^{-1} - I|| \to_{a.s.} 0$.

Definition 6.2.
$\mathbb{V}_n \equiv \sqrt{n}(\mathbb{G}_n^{-1} - I) \equiv$ the uniform quantile process.
$\sqrt{n}(\mathbb{F}_n^{-1} - F^{-1}) \equiv$ the quantile process $\stackrel{d}{=} \sqrt{n}(F^{-1}(\mathbb{G}_n^{-1}) - F^{-1})$.

Theorem 6.3.

$$\mathbb{V}_n = -\mathbb{U}_n(\mathbb{G}_n^{-1}) + \sqrt{n}(\mathbb{G}_n \circ \mathbb{G}_n^{-1} - I),$$

and hence for the specially constructed $\mathbb{U}_n$ of theorem 4.2 it follows that, with $\mathbb{V} \equiv -\mathbb{U} \overset{d}{=} \mathbb{U}$,

$$||\mathbb{V}_n - \mathbb{V}|| \rightarrow_{a.s.} 0 \quad \text{as} \quad n \rightarrow \infty.$$

Proof.

$$\begin{aligned}
\mathbb{V}_n &\equiv \sqrt{n}(\mathbb{G}_n^{-1} - I) \\
&= \sqrt{n}(\mathbb{G}_n^{-1} - \mathbb{G}_n(\mathbb{G}_n^{-1})) + \sqrt{n}(\mathbb{G}_n(\mathbb{G}_n^{-1}) - I) \\
&= -\mathbb{U}_n(\mathbb{G}_n^{-1}) + \sqrt{n}(\mathbb{G}_n \circ \mathbb{G}_n^{-1} - I).
\end{aligned}$$

Now $||\mathbb{G}_n^{-1} - I|| \rightarrow_{a.s.} 0$ by proposition 6.1, and

$$||\mathbb{G}_n \circ \mathbb{G}_n^{-1} - I|| = \sup_{0 \le t \le 1} ||\mathbb{G}_n(\mathbb{G}_n^{-1}(t)) - t| = \frac{1}{n}.$$

Hence

$$\begin{aligned}
||\mathbb{V}_n - \mathbb{V}|| &\le ||\mathbb{U}_n(\mathbb{G}_n^{-1}) - \mathbb{U}|| + \frac{1}{\sqrt{n}} \\
&\le ||\mathbb{U}_n(\mathbb{G}_n^{-1}) - \mathbb{U}(\mathbb{G}_n^{-1})|| + ||\mathbb{U}(\mathbb{G}_n^{-1}) - \mathbb{U}|| + \frac{1}{\sqrt{n}} \\
&\le ||\mathbb{U}_n - \mathbb{U}|| + ||\mathbb{U}(\mathbb{G}_n^{-1}) - \mathbb{U}|| + \frac{1}{\sqrt{n}} \\
&\rightarrow_{a.s.} 0 + 0 + 0 = 0
\end{aligned}$$

since $\mathbb{U}$ is a continuous (and hence uniformly continuous) function on $[0,1]$. □

Theorem 6.4. Let $Q = F^{-1}$, and suppose that Q is differentiable at $0 < t_1 < \cdots < t_k < 1$. Then

$$\begin{pmatrix} \sqrt{n}(\mathbb{F}_n^{-1}(t_1) - F^{-1}(t_1)) \\ \cdot \\ \cdot \\ \cdot \\ \sqrt{n}(\mathbb{F}_n^{-1}(t_k) - F^{-1}(t_k)) \end{pmatrix} \rightarrow_d \begin{pmatrix} Q'(t_1)\mathbb{V}(t_1) \\ \cdot \\ \cdot \\ \cdot \\ Q'(t_k)\mathbb{V}(t_k) \end{pmatrix} \sim N_k(0, \Sigma)$$

where

$$\Sigma \equiv (\sigma_{ij}) = (Q'(t_i)Q'(t_j)(t_i \wedge t_j - t_i t_j)).$$

Moreover, if Q' is nonzero and continuous on $[a,b] \subset [0,1]$, then for any $[c,d] \subset [a,b]$

$$\|\sqrt{n}(F^{-1}(\mathbb{G}_n^{-1}) - F^{-1}) - Q'\mathbb{V}\|_c^d \rightarrow_{a.s.} 0 \quad \text{as} \quad n \rightarrow \infty .$$

Note that $Q'(t) = 1/f(F^{-1}(t))$.

Proof. Suppose that $k = 1$ and let $t_1 \equiv t$. Then

$$\begin{aligned}
\sqrt{n}(\mathbb{F}_n^{-1}(t) - F^{-1}(t)) &\stackrel{d}{=} \sqrt{n}(Q(\mathbb{G}_n^{-1}(t)) - Q(t)) \\
&= \frac{Q(\mathbb{G}_n^{-1}(t)) - Q(t)}{\mathbb{G}_n^{-1}(t) - t} \sqrt{n}(\mathbb{G}_n^{-1}(t) - t) \\
&\stackrel{d}{=} \frac{Q(\mathbb{G}_n^{-1}(t)) - Q(t)}{\mathbb{G}_n^{-1}(t) - t} \mathbb{V}_n(t) \qquad \text{for the special } \mathbb{V}_n \text{ process} \\
&\rightarrow_{a.s.} Q'(t)\,\mathbb{V}(t) \sim N(0, (Q'(t))^2\, t(1-t)) .
\end{aligned}$$

Similarly,

$$\begin{pmatrix} \sqrt{n}(\mathbb{F}_n^{-1}(t_1) - F^{-1}(t_1)) \\ \cdot \\ \cdot \\ \cdot \\ \sqrt{n}(\mathbb{F}_n^{-1}(t_k) - F^{-1}(t_k)) \end{pmatrix} \stackrel{d}{=} \sqrt{n} \begin{pmatrix} Q(\mathbb{G}_n^{-1}(t_1) - Q(t_1) \\ \cdot \\ \cdot \\ \cdot \\ Q(\mathbb{G}_n^{-1}(t_1) - Q(t_1) \end{pmatrix} \rightarrow_{a.s.} \begin{pmatrix} Q'(t_1)\mathbb{V}(t_1) \\ \cdot \\ \cdot \\ \cdot \\ Q'(t_1)\mathbb{V}(t_k) \end{pmatrix} .$$

Corollary 6.5. (Asymptotic normality of the median). Suppose that $Q = F^{-1}$ is differentiable at $t = 1/2$. Then

$$\sqrt{n}(\mathbb{F}_n^{-1}(\frac{1}{2}) - F^{-1}(\frac{1}{2})) \rightarrow_d Q'(\frac{1}{2})\, N(0, \frac{1}{4}) = N(0, \frac{1/4}{f^2(F^{-1}(1/2))}) .$$

Corollary 6.6. Suppose that J is bounded and continuous a.e. F^{-1} , and suppose that $E(X^2) < \infty$. Let

$$T_n \equiv \frac{1}{n}\sum_{i=1}^{n} J(\frac{i}{n+1})\, X_{(i)} , \qquad \mu = \int_0^1 J(u)F^{-1}(u)\, du .$$

Then

$$\sqrt{n}(T_n - \mu) \rightarrow_d \int_0^1 J\, \mathbb{V}\, dF^{-1} \sim N(0, \sigma^2(J, F))$$

where

$$\sigma^2(J, F) = \int_0^1 \int_0^1 J(s)J(t)\,(s \wedge t - st)\, dF^{-1}(s)\, dF^{-1}(t) .$$