§5.1 BASIC CONCEPTS of Random Samples

Def 5.1.1 A $\boxed{\text{random sample}}$, $X_1, X_2, \dots, X_n$, size n from $f(x)$ if

- $X_1, X_2, \dots X_n$ are mutually independent
- independent } Note if discrete must be sampled w/ replacement
- identical

} iid

- If we only have one observation, $n=1$, we can calculate probabilities using $f_X(x)$
- otherwise, $n>1$, $f_X(x_1, x_2, \dots x_n) = f_X(x_1) f_X(x_2) \cdots f_X(x_n) = \prod_1^n f_X(x_i)$ As X_i are iid

Example 5.1.2. Let $X_1, X_2, \dots X_n$ be a random sample from an exponential(β)

ie $X_1, X_2, \dots X_n$ correspond to the times until failure

- $f_X(x_1, x_2, \dots x_n | \beta) = \prod_1^n f_X(x_i|\beta) = \prod_1^n \frac{1}{\beta} e^{-x_i/\beta}$
$= \frac{1}{\beta^n} e^{-1/\beta \sum x_i}$

- What is the probability that all boards last more than two years

$= P(X_1>2, X_2>2, X_3>2, \dots X_n>2)$

$= \int_2^\infty \int_2^\infty \cdots \int_2^\infty f_X(x_1, x_2, \dots x_n)\, dx_1 dx_2 \dots dx_n$

$= e^{-2n/\beta}$ • If β, the avg life of a board is large, in comparison, to n this approaches 1

Simple random sample - Sampling w/o replacement from a finite population

- We can use what we know from 5.1.1 in only a special case
  - If $\frac{1}{\text{Population size} - \text{sample size}} \approx \frac{1}{\text{Pop. size}}$

example: $\{1, 2, \dots 1000\}$ = finite population

$n = 10 \longrightarrow \frac{1}{990} \approx \frac{1}{1000}$

Approx w/ 5.1.1 $P(X_1>200, X_2>200, \dots X_{10}>200) = \prod_1^{10}(X_i>200) = \left(\frac{800}{1000}\right)^{10} = .107374$

Actual $Y = \#$ in sample >200 $Y \sim$ hypergeometric$(N=1000, M=800, K=10)$

$P(Y=10) = \dfrac{\binom{800}{10}\binom{200}{0}}{\binom{1000}{10}} = .106164$

We Note they are close

"Burn Burn like a star"

§5.2

Def 5.2.1 • Let $X_1, X_2, \ldots X_n$ be a random (iid) sample of size n from a population and let $T(x_1, x_2, \ldots x_n)$ be a real or vector-valued function whose domain includes the sample space $(X_1, X_2, \ldots, X_n)$. Then the Random Variable $Y = T(X_1, X_2, \ldots X_n)$ is called a statistic. The probability distribution of a statistic is called the sampling distribution

- A statistic musn't depend on a parameter

Def 5.2.2 · Sample mean $= \bar{x} = \frac{x_1 + x_2 \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Def 5.2.3 · $S^2 = \underbrace{\frac{1}{n-1}}_{\text{to make it unbiased}} \sum_{i=1}^{n}(x_i - \bar{x})^2$ Sample Variance

Note: $S = \sqrt{S^2}$ Sample st deviation

Theorem 5.2.4: Let $x_1, x_2, \ldots x_n$ be any numbers and $\bar{x} = \frac{(x_1 + x_2 + \cdots x_n)}{n}$ then
Proof p. 212
a) $\min \sum_i (x_i - a)^2 = \sum_i (x_i - \bar{x})^2$
b) $(n-1)S^2 = \sum_i (x_i - \bar{x})^2 = \sum_i x_i^2 - n\bar{x}^2$

Lemma 5.2.5 Let $X_1, X_2, \ldots X_n$ be a random (iid) sample from a population &
Proof p. 213 let $g(x)$ be a function $\ni E(g(X_1))$ and $var(g(X))$ exist then
a) $E(\sum_i g(X_i)) = nE(g(X_1))$
b) $var(\sum_i g(X_i)) = n(var(g(X_1))$

Theorem 5.2.6 Let $X_1, X_2, \ldots X_n$ be a random sample from a population w/ mean
p 214 μ and $\sigma^2 < \infty$ Then
a) $E(\bar{x}) = \mu \rightarrow$ unbiased estimator of μ
b) $Var(\bar{x}) = \sigma^2/n$
c) $E(S^2) = \sigma^2 \rightarrow$ unbiased estimator of σ^2

Acoustic You are my sunshine

→ Since $\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots X_n)$

- thus $f_{\bar{x}}(x) = n f_x(nx)$
- thus $M_{\bar{x}}(t) = E(e^{t\bar{x}}) = E(e^{t(x_1+x_2+\cdots x_n)/n}) = E(e^{(t/n)Y}) = M_Y(t/n)$

Theorem 5.2.7: Let $X_1, X_2, \ldots X_n$ be a random sample from a population w/ MGF $M_X(t)$. Then the MGF of the sample mean is:

$$M_{\bar{x}}(t) = [M_X(t/n)]^n$$

<u>Example</u> Let $X_1, X_2, \ldots, X_n$ be a random sample from a $n(\mu, \sigma^2)$ population. The MGF of the sample mean is

$$M_{\bar{x}}(t) = [M_X(t/n)]^n = \left[e^{\mu\frac{t}{n} + \frac{\sigma^2(t/n)^2}{2}}\right]^n = e^{\mu t + \frac{\sigma^2 t^2}{n}/2}$$

Thus $\bar{x} \sim N(\mu, \sigma^2/n)$

Theorem 5.2.9 If X and Y are independent continuous RV w/ PDFs f_x, f_y

Proof p215

- then pdf of $Z = X+Y$ is
  $$f_Z(Z) = \int_{-\infty}^{\infty} f_x(w) f_y(z-w)\,dw$$
- then pdf of $Z = X-Y$
- then pdf of $Z = XY$
- then pdf of $Z = X/Y$

Example of Convolution

$Z_1, Z_2, \ldots Z_n$ iid Cauchy (0,1)

Let $U \sim$ Cauchy$(0, \sigma)$ $V \sim$ Cauchy$(0, \tau)$

$f_U(u) = \frac{1}{\pi\sigma} \cdot \frac{1}{1+(u/\sigma)^2}$ $f_V(v) = \frac{1}{\pi\tau} \cdot \frac{1}{1+(v/\tau)^2}$ $-\infty < u, v < \infty$

By Theorem 5.2.9:

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\pi(\sigma+\tau)} \quad \frac{1}{1+(z/(\sigma+\tau))^2} \quad -\infty < z < \infty \qquad Z = U+V$$

∴ [Cauchy $(0,\tau)$ + Cauchy $(0,\sigma)$] $\sim$ Cauchy $(0, \tau+\sigma)$

- Suppose $X_1, X_2, \ldots X_n$ is a random sample from $\frac{1}{\sigma} f((x-\mu)/\sigma)$ a location scale family
- From theorem 3.5.6 for a location scale family $\exists$ random variables $Z_1, Z_2, \ldots Z_n \ni X_i = \sigma Z_i + \mu$ and the pdf of each Z_i is $f_Z(z)$. Furthermore $Z_1, Z_2, \ldots Z_n$ are mutually independent thus $Z_1, Z_2, \ldots Z_n$ is a random sample from $f(z) \ni$

$$\bar{X} = \frac{1}{n}\sum_1^n X_i = \frac{1}{n}\sum_1^n(\sigma Z_i + \mu) = \frac{1}{n}(\sigma \sum_1^n Z_i + n\mu) = \sigma \bar{Z} + \mu$$

- Also from theorem 3.5.6 we find that if $g(z)$ is the pdf of $\bar{Z}$ then $(\frac{1}{\sigma}) g((x-\mu)/\sigma)$ is the pdf of $\bar{X}$

Theorem 5.2.11:
- Suppose $X_1, X_2, \ldots X_n$ is a random sample from pdf or pmf $f(x|\theta)$ where $f(x|\theta) = h(x)\, c(\theta) \exp(\sum_1^k w_i(\theta) t_i(x))$ is a member of an exp. family
- Define statistics $T_1, T_2, \ldots T_k$ by $T_i(X_1, X_2, \ldots X_n) = \sum_{j=1}^n t_i(X_j)$ $i = 1,2,3,\ldots k$
- If the set $\{(w_1(\theta), w_2(\theta), \ldots w_k(\theta)), \theta \in \Theta\}$ contains an open subset of $\mathbb{R}^k$, then the dist of $(T_1, T_2, \ldots T_k)$ is an exp family of form

$$f_T(u_1, u_2, \ldots u_k|\theta) = H(u_1, u_2, \ldots u_k)\,[c(\theta)]^n \exp(\sum_1^k w_i(\theta) u_i)$$

example: Suppose $X_1, X_2, \ldots X_n$ is a random sample from a Bernoulli(p) dist. w/ $n=1$, Bernoulli(p) is an exponential family w/ $k=1$, $c(p) = (1-p)$, $w_1(p) = \log(p/(1-p))$ and $t_1(x) = x$. Thus by 5.5.11

$$T_1 = T_1(X_1, X_2, \ldots X_n) = \sum_1^n X_i \quad \& \quad T_1 \sim \text{bin}(n,p)$$

NOTE: Binomial(n,p) is an exponential family w/ <u>the same</u> $w_1(p)$, $c(p)$

§5.3

Theorem 5.3.1 Let $X_1, X_2, \ldots X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution and let $\bar{X} = \frac{1}{n}\sum_1^n X_i$ and $S^2 = \frac{1}{(n-1)}\sum(X_i - \bar{X})^2$ then

a) $\bar{X}$ and S^2 are independent

b) $\bar{X}$ has $N(\mu, \sigma^2/n)$ distribution

c) $(n-1)S^2/\sigma^2$ has chi-squared distribution w/ n-1 degrees of freedom

Proof p. 218

Lemma 5.3.2 Chi-squared(p) $\equiv \chi^2_{(p)}$

a) If Z is $N(0,1)$ then $Z^2 \sim \chi^2_{(1)}$

b) If $X_1, X_2, \ldots X_n$ are independent $\chi^2_{(p)}$ then $\sum_1^n X_i \sim \chi^2_{(np)}$

Lemma 5.3.3 Let $X_j \sim N(\mu_j, \sigma_j^2)$, $j = 1, 2, \ldots n$, independent. For constants a_{ij} and b_{rj} ($j = 1, 2, \ldots n$, $i = 1, 2, \ldots k$ and $r = 1, 2, \ldots m$) where $k + m \leq n$

$U_i = \sum_{j=1}^n a_{ij} X_j \quad i = 1, 2, \ldots k \qquad V_r = \sum_{j=1}^n b_{rj} X_j \quad r = 1, 2, \ldots m$

a) U_i and V_r are independent iff $\operatorname{Cov}(U_i, V_r) = 0$ futhermore, $\operatorname{Cov}(U_i, V_r) = \sum_{j=1}^n a_{ij} b_{rj} \sigma_j^2$

b) The random vectors $(U_1, \ldots U_k)$ and $(V_1, \ldots V_m)$ are independent iff U_i is independent of V_r $\forall$ pairs i, r ($i = 1, 2, \ldots k$; $r = 1, 2, \ldots, m$

* If $X_1, X_2, \ldots X_n$ is from $N(\mu, \sigma^2)$ then $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

Def 5.3.4 Let $X_1, X_2, \ldots X_n$ be a random sample from $N(\mu, \sigma^2)$ then

$\left(\frac{\bar{X} - \mu}{S/\sqrt{n}}\right) \sim t_{(n-1)} \qquad f_T(t) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma(p/2)} \frac{1}{(p\pi)^{1/2}} \frac{1}{(1 + t^2/p)^{(p+1)/2}}$

* $t_{(1)} \equiv$ Cauchy
* T distribution has no MGF – only produces deg. of freedom - 1 moments

Example · Let $X_1, X_2, \ldots X_n$ be a random sample from a $N(\mu_x, \sigma_x^2)$ population and let $Y_1, Y_2 \ldots Y_m$ be a random sample from $N(\mu_y, \sigma_y^2)$

· We're interested in σ_x^2 / σ_y^2

· Information about this ratio is contained in S_x^2 / S_y^2

· F can give us the distribution of

$$\frac{S_x^2 / S_y^2}{\sigma_x^2 / \sigma_y^2} = \frac{S_x^2 / \sigma_x^2}{S_y^2 / \sigma_y^2}$$

as S_x^2/σ_x^2 and S_y^2/σ_y^2 are each scaled χ^2 and indep.

Def 5.3.6 From above, $F = \dfrac{S_x^2 / \sigma_x^2}{S_y^2 / \sigma_y^2} \sim F_{(n-1, m-1)}$

for dof p, q

$$f_F(f) = \frac{\Gamma\left(\frac{p+q}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)} \left(\frac{p}{q}\right)^{p/2} \frac{x^{(p/2)-1}}{\left[1+(p/q)x\right]^{p+q/2}} \quad 0 < x < \infty$$

Cont Example As $(S_x^2/\sigma_x^2)/(S_y^2/\sigma_y^2) \sim F_{(n-1, m-1)}$

we consider

$$E(F_{(n-1,m-1)}) = E\left(\frac{\chi^2_{n-1}/(n-1)}{\chi^2_{m-1}/(m-1)}\right) = E\left(\frac{\chi^2_{n-1}}{n-1}\right) E\left(\frac{\chi^2_{m-1}}{m-1}\right)$$

$$= \left(\frac{n-1}{n-1}\right)\left(\frac{m-1}{m-3}\right) = \frac{m-1}{m-3} \approx 1 \text{ for Large } m$$

NOTE: only finite & pos. for $m > 3$

Theorem 5.3.8

a) If $X \sim F_{p,q}$ then $\frac{1}{X} \sim F_{q,p}$

b) If $X \sim t_q$ then $X^2 \sim F_{1,q}$

c) If $X \sim F_{p,q}$ then $(p/q)X / (1 + (p/q)X) \sim \text{beta}(p/2, q/2)$

5.4 Order statistics

Def 5.4.1 $X_1, X_2, \ldots X_n$ are the sample values
$X_{(1)}, X_{(2)}, X_{(3)}, \ldots X_{(n)}$, the values listed in ascending order are called <u>ordered statistics</u>

i.e $X_{(1)} \le X_{(2)} \le X_{(3)} \le \cdots \le X_{(n)}$

$X_{(1)} = \min(X_i)$

$X_{(n)} = \max(X_i)$

* Since these are random variables we can discuss the probability of them holding certain values

ie Range $= X_{(n)} - X_{(1)}$

$$\text{Median} = \begin{cases} X_{((n+1)/2)} & n \text{ even} \\ (X_{(n/2)} + X_{(n/2+1)})/2 & n \text{ odd} \end{cases}$$

* Percentile: $\forall$ # $p \in [0,1]$ the $(100*p)^{th}$ percentile is the observation $\ni$ approximately np of the observations are less than this and $n(1-p)$ are greater

Theorem 5.4.3 · Let $X_1, X_2, \ldots X_n$ be a random sample from a discrete dist. w/ pmf $f_X(x_i) = p_i$ where $x_1 < x_2 < \ldots$ are the possible values of X in ascending order.

* Define: $P_0 = 0$, $P_1 = p_1$, $P_2 = p_1 + p_2$, $\vdots$, $P_i = p_1 + p_2 + \cdots + p_i$

* Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denote the orderstats of the sample

THEN

$$P(X_{(j)} \le x_i) = \sum_{k=j}^{n} \binom{n}{k} P_i^k (1-P_i)^{n-k}$$

and

$$P(X_{(j)} = x_i) = \sum_{k=j}^{n} \binom{n}{k} \left[P_i^k (1-P_i)^{n-k} - P_{i-1}^k (1-P_{i-1})^{n-k} \right]$$

Proof 228

Theorem 5.4.4 Let $X_{(1)}, X_{(2)}, \ldots X_{(n)}$ denote the order statistics of a random sample from a continuous population w/ CDF $F_X(x)$ and pdf $f_X(x)$ Then the pdf of $X_{(j)}$ is

$$* \quad f_{X_{(j)}}(x) = \frac{n!}{(j-1)!\,(n-j)!} f_X(x) \left[F_X(x)\right]^{j-1} \left[1 - F_X(x)\right]^{n-j}$$

Proof p. 229

<u>Example 5.4.5</u> Let $X_1, X_2, \ldots X_n$ be iid uniform $(0,1)$ so $f_X(x) = 1$ for $x \in (0,1)$ and $F_X(x) = x$ for $x \in (0,1)$.

$$f_X(x) = 1 \; I(0<x<1)$$

$$F_X(x) = {}_0\!\int^1 1\,dx = x\Big|_0^1 = x$$

$$f_{X(j)}(x) = \frac{n!}{(j-1)!\,(n-j)!} (1) \left[x\right]^{j-1} \left[1-x\right]^{n-j}$$

$$= \frac{n!}{(j-1)!\,(n-j)!} \left[x\right]^{j-1} \left[1-x\right]^{n-j}$$

NOTE: we can rewrite this $\Gamma(\alpha) = (\alpha-1)!$

$$= \frac{\Gamma(n+1)}{\Gamma(j)\,\Gamma(n-j+1)} x^{j-1} (1-x)^{(n-j+1)-1}$$

So $E(X_{(j)}) = \frac{j}{n+1}$ $\quad var(X_{(j)}) = \frac{j(n-j+1)}{(n+1)^2(n+2)}$

<u>Theorem 5.4.6</u> Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be ordered statistics of a continuous population w/ CDF $F_X(x)$ and pdf $f_X(x)$. $1 < i < j < n$

$$f_{x(i),x(j)}(u,v) = \frac{n!}{(i-1)!\,(j-1-i)!\,(n-j)!} f_X(u) f_X(v) \left[F_X(u)\right]^{i-1}$$

$$* \left[F_X(v) - F_X(u)\right]^{j-1-i} \left[1 - F_X(v)\right]^{n-j}$$

Example 5.47 Let $X_1, X_2, \ldots X_n \overset{iid}{\sim}$ uniform$(0,a)$ and Let $X_{(1)}, X_{(2)}, \ldots X_{(n)}$ be the ordered statistics $R = X_{(n)} - X_{(1)}$ and $R_{mid} = \frac{X_{(1)} + X_{(n)}}{2} = V$

Joint dist

$$f_{X_{(1)}X_{(n)}}(x_1, x_n) = \frac{n(n-1)}{a^2}\left(\frac{x_n}{a} - \frac{x_1}{a}\right)^{n-2} \quad 0 < x_1 < x_n < a$$

$$= \frac{n(n-1)(x_n - x_1)^{n-2}}{a^n} \quad 0 < x_1 < x_n < a$$

$$X_{(1)} = V - R/2 \quad X_{(n)} = V + R/2$$

$$|J| = \begin{vmatrix} \frac{dx_{(1)}}{dv} & \frac{dx_{(1)}}{dR} \\ \frac{dx_{(n)}}{dv} & \frac{dx_{(n)}}{dR} \end{vmatrix} = \begin{vmatrix} 1 & -1/2 \\ 1 & 1/2 \end{vmatrix} = |1/2 - -1/2| = 1$$

$$\therefore f_{R,V}(r,v) = \frac{n(n-1)\left((v + R/2) - (v - R/2)\right)^{n-2}}{a^n} \quad \begin{pmatrix} 0 < r < a \\ r/2 < v < a - r/2 \end{pmatrix}$$

$$= \frac{n(n-1)(r)^{n-2}}{a^n}$$

$$f_R = \int_{r/2}^{a-r/2} \frac{n(n-1)(r)^{n-2}}{a^n}\, dv = -\frac{n(n-1)(r)^{n-2}\left(\frac{r}{2}\right)}{a^n} + \frac{n(n-1)(r)^{n-2}(a - r/2)}{a^n}$$

$$= \frac{n(n-1)r^{n-2}(a-r)}{a^n}$$

↑ Distribution of $X_{(n)} - X_{(1)}$

5.5 CONVERGENCE   look at $\bar{X}_n$ as $n \to \infty$

Theorem 5.5.2 **Weak Law of Large Numbers (WLLN)**

- Let $X_1, X_2, \ldots$ be iid random variables w/ $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$

Define $\bar{X}_n = \frac{1}{n}\sum_1^n X_i$ Then $\forall\ \epsilon > 0$

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| < \epsilon) = 1 \quad \equiv \bar{X}_n \text{ converges in probability to } \mu$$

Proof p.233

Example: $X_1, X_2, \ldots$ of iid RV w/ $E(X) = \mu$ and $Var(X) = \sigma^2$

Define $S^2 = \frac{1}{n-1}\sum_1^n (x_i - \bar{x}_n)^2$

By Chebychev $P(|S_n^2 - \sigma^2| \geq \epsilon) \leq \frac{E(S_n^2 - \sigma^2)^2}{\epsilon^2} = \frac{var(S_n^2)}{\epsilon^2}$

- If $\frac{var(S_n^2)}{\epsilon^2} \to 0 \quad S_n^2 \to \sigma^2$ (in probability)

Theorem 5.5.4 Suppose that $X_1, X_2, \ldots$ converges in probability to random variable X and that h is a continuous function. Then $h(X_1), h(X_2), \ldots$ converge in probability to $h(X)$ by continuity

Example 5.5.5 If S_n^2 is a consistent estimator of σ^2, then by theorem 5.5.4 the sample standard deviation $S_n = \sqrt{S_n^2} = h(S_n^2)$ is a consistent estimator for σ

- as $0 < S_n^2 < \infty \quad h(x) = \sqrt{x}$ is continuous 1-1 on $(0, \infty)$

Def 5.5.1 Converges in Probability $\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0 \quad \equiv \lim_{n\to\infty} P(|X_n - X| \leq \epsilon) = 1$

*Example for next definition

Convergence almost surely

- Let S be the sample space $[0,1]$ w/ the uniform dist.
- Define $X_n(s) = s + s^n$ and $X(s) = s$
- $\forall\ s \in [0,1)$ $s^n \to 0$ as $n \to \infty$ and $X_n(s) \to s = X(s)$ However $X_n(1) = 2\ \forall n$

so $X_n(1)$ doesn't converge to $1 = X(1)$

but since $X_n(s) \to s = X(s)$ on $[0,1)$ we say $P([0,1)) = 1$, $X_n \to X$ almost surely

Def 5.5.6 A sequence of RV $X_1, X_2, \dots$ converges almost surely to a random variable X if $\forall\ \epsilon > 0$

$$P(\lim_{n\to\infty} |X_n - X| < \epsilon) = 1$$

* Almost surely → in probability

*Ex 5.5.8 (Convergence in probability but not almost surely)

Tobs for help

- Sample Space $= S = [0,1]$ w/ uniform probability
- Define $X_1, X_2, \dots$

$$X_1(s) = s + I_{[0,1]}(s)$$
$$X_2(s) = s + I_{[0,1/2]}(s)$$
$$X_3(s) = s + I_{[1/2,1]}(s)$$
$$X_4(s) = s + I_{[0,1/3]}(s)$$
$$X_5(s) = s + I_{[\frac{1}{3},\frac{2}{3}]}(s)$$
$$X_6(s) = s + I_{[\frac{2}{3},1]}(s)$$

- Let $X(s) = s$ As $n \to \infty$ $P(|X_n - X| \geq \epsilon)$

Theorem 5.5.9 SLLN

- Let $X_1, X_2, \ldots$ be iid random variables w/ $E(X_i) = \mu$ and $var(X_i) = \sigma^2 < \infty$ and define $\bar{X}_n = (1/n)\sum_i^n X_i$ then $\forall\ \epsilon > 0$

$$P(\lim_{n\to\infty} |\bar{X}_n - \mu| < \epsilon) = 1$$

that is, $\bar{X}_n$ converges almost surely to μ

~~Def 5.5.10~~ A sequence of random variables $X_1, X_2, \ldots$ converge in distribution if $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ at all points x where $F_X(x)$ is continuous

Example 5.5.11 If $X_1, X_2, \ldots$, are iid uniform$(0,1)$ and $X_{(n)} = \max(X_i)$, Let us examine where $X_{(n)}$ converges in distribution. As $n \to \infty$ we expect $X_{(n)}$ to get close to 1 and as $X_{(n)}$ must be less than 1 we have $\forall\ \epsilon > 0$

$$P(|X_{(n)} - 1| \geq \epsilon) = P(X_{(n)} \geq 1 + \epsilon) + P(X_{(n)} \leq 1 - \epsilon)$$

$$= 0 + P(X_{(n)} \leq 1 - \epsilon) \quad \text{as } X \sim U(0,1)$$

$$= P(X_i < 1 - \epsilon \text{ for } i = 1, \ldots, n)$$

$$= \prod_{i}^{n} P(X_i < 1 - \epsilon) = (1 - \epsilon)^n \text{ which goes to zero}$$

$\therefore X_{(n)}$ converges in probability to one

* Take $\epsilon = t/n$

$$P(X_i < 1 - t/n) = (1 - t/n)^n \to e^{-t} \text{ as } n \to \infty$$

$$P(t \geq n(X_i - 1)) \to 1 - e^{-t}$$

$\therefore n(X_i - 1)$ converges in distribution to $\exp(1)$

Theorem 5.5.12 If the sequence of random variables $X_1, X_2, X_3 \ldots$ converges in probability to a RV X the sequence also converges in distribution to X

$\Updownarrow$

Theorem 5.5.13 The sequence of random variables $X_1, X_2, \ldots$ converges in probability to a constant μ iff the sequence also converges in distribution to μ

ie $P(|X_n - X| \geq \epsilon) \to 0 \quad \forall\ \epsilon > 0$

$\equiv$

$$P(X_n \leq x) \to \begin{cases} 0 & (x < \mu) \\ 1 & (x \geq \mu) \end{cases}$$

Theorem 5.5.14

Central Limit Theorem: • Let $X_1, X_2, \ldots$ be a seq of iid RV ~~whose mgfs exist in a neighborhood of 0.~~

not needed → ~~• that is $M_X(t)$ exists for $|t|<h$ for some positive h~~

• Let $E(X_i)=\mu$ and $Var(X_i)=\sigma^2$

• Define $\bar{X}_n = \frac{1}{n}\sum_i^n X_i$

• Let $G_n(x)$ denote the CDF of $\sqrt{n}(\bar{X}_n-\mu)/\sigma$ then $\forall x \ni -\infty < x < \infty$

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

$\therefore \sqrt{n}(\bar{X}_n-\mu)/\sigma \sim N(0,1)$

Proof p 237

Example 5.5.16 Suppose $X_1, \ldots, X_n$ are a random sample from a negative binomial (r,p) distribution

$$E(X) = \frac{r(1-p)}{p} \qquad Var(X) = \frac{r(1-p)}{p^2}$$

The Central Limit theorem says

$$\frac{\sqrt{n}\,(\bar{X} - r(1-p)/p)}{\sqrt{r(1-p)/p^2}} \sim N(0,1)$$

Theorem 5.5.17

Slutsky's Theorem If $X_n \to X$ in distribution and $Y_n \to a$, a constant, in probability then

a) $Y_n X_n \to aX$ in distribution

b) $X_n + Y_n \to X + a$ in distribution

Example 5.5.18 Suppose $\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma} \to N(0,1)$

• As before $\lim_{n\to\infty} S_n^2 = 0$ then $S_n^2 \to \sigma^2$ in probability

• Ex 5.5.3: $\sigma/S_n \to 1$ in probability

By Slutsky's

$$\frac{\sigma/S_n \,\sqrt{n}(\bar{X}_n-\mu)}{\sigma} \to (1)\, N(0,1)$$

For $X_1, X_2, \ldots X_n$ Bernoulli(p) trials

odds $= p/1-p$ "has odds $p/1-p$ of success"

odds ratio $= (p/1-p)/(\frac{r}{1-r})$ - gives relative odds" of one success over the other

Def 5.5.20 If function $g(x)$ has derivatives of order r, that is, $g^{(r)}(x) = \frac{d^r}{dx^r} g(x)$ exists, then for any constant a, <u>the Taylor polynomial of order r about a</u> is

$$T_r(x) = \sum_{i=0}^{r} \frac{g^{(i)}(a)}{i!}(x-a)^i$$

Theorem 5.5.21 If $g^{(r)}(a) = \frac{d}{dx^r} g(x)|_{x=a}$ exists then $\lim_{x \to a} \frac{g(x) - T_r(x)}{(x-a)^r} = 0$

<u>Example 5.5.22</u> • We are interested in $\frac{\hat{p}}{1-\hat{p}}$ as an estimate of $\frac{p}{1-p}$ where p is a binomial success probability

• Take $g(p) = \frac{p}{1-p}$ so $g'(p) = \frac{1}{(1-p)^2}$ and

$$\operatorname{var}\left(\frac{\hat{p}}{1-\hat{p}}\right) \approx [g'(p)]^2 \operatorname{Var}(\hat{p})$$

$$= \left[\frac{1}{(1-p)^2}\right]^2 \frac{p(1-p)}{n} = \frac{p}{n(1-p)^3}$$

* Note for RV $T_1, T_2, \ldots T_k$ w/ means $\theta_1, \theta_2, \ldots \theta_k$ $\exists$ a differentiable function $g(\underset{\sim}{T})$ $\ni$

$$E_{\underset{\sim}{\theta}}(g(\underset{\sim}{T})) \approx g(\underset{\sim}{\theta}) \quad \& \quad Var(g(\underset{\sim}{T})) \approx [g'(\underset{\sim}{\theta})]^2 Var(\underset{\sim}{T})$$

<u>Example 5.5.23</u> X is RV $E_\mu(x) = \mu \neq 0$

• an estimate of $g(\mu)$: $g(X) = g(\mu) + g'(\mu)(x-\mu)$

$\therefore E(g(X)) \approx g(\mu)$ & $\operatorname{var}(g(X)) \approx [g'(\mu)]^2 Var_\mu(x)$

Theorem 5.5.24

Delta Method Let Y_n be a sequence of RV that satisfies $\sqrt{n}(Y_n - \theta) \to n(0, \sigma^2)$ in distribution. For a given function g and a specific value θ suppose that $g'(\theta)$ exists and is not 0 then

$$\sqrt{n}[g(Y_n) - g(\theta)] \to N(0, \sigma^2[g'(\theta)]^2) \text{ in distribution}$$

Proof 243

Example 5.5.25 $\bar{X}$ mean of a random sample

$E(X) = \mu \neq 0$

$E(g(x)) \approx g(\mu)$ $\qquad var(g(x)) \approx [g'(\mu)]^2 Var_\mu(x)$

ie $g(\mu) = 1/\mu$

$E(g(x)) \approx \frac{1}{\mu}$ $\qquad var(g(x)) \approx (\frac{1}{\mu^2})^2 var_\mu(X)$

By Delta method

$$\sqrt{n}\left(\frac{1}{\bar{x}} - \frac{1}{\mu}\right) \to N\left(0, \frac{1}{\mu^4} Var_\mu(X_1)\right)$$

Theorem 5.5.26

Second order deltamethod Let Y_n be a sequence of random variables that satisfies $\sqrt{n}(Y_n - \theta) \to N(0, \sigma^2)$ in distribution. For a given function g and a specific value of θ suppose that $g'(\theta) = 0$ and $g'(\theta) = 0$ and $g''(\theta) \neq 0$

Then:

$$n[g(Y_n) - g(\theta)] \to \sigma^2 \frac{g''(\theta)}{2} \chi^2_{(1)} \text{ in distribution}$$

Theorem 5.5.28

Let $X_1, X_2, \ldots X_n$ be a random sample w/ $E(X_{ij}) = \mu_i$ and $cov(X_{ik}, X_{jk}) = \sigma_{ij}$. For a given function g w/ continuous first partial derivatives and a specific value of $\mu = (\mu_1, \mu_2 \ldots \mu_n)$ for which $\tau^2 = \sum\sum \sigma_{ij} \frac{\partial g(\mu)}{\partial \mu_i} \cdot \frac{\partial g(\mu)}{\partial \mu_j} > 0$

$$\sqrt{n}[g(\bar{X}_1, \ldots \bar{X}_s) - g(\mu_1, \ldots, \mu_p)] \to N(0, \tau^2) \text{ in distribution}$$