STA 453/553 CLASS 1

- INTRODUCE CLASS AND SYLLABUS DISCUSSION

WEB SITE: HTTP://WWW.STAT.UNH.EDU/~GHUERTA/INFERENCE/COURSE.HTML.

BEGIN WITH CHAP. 5: ELEMENTS OF STAT. INFERENCE.

X A R.V. THAT FOLLOWS SOME DIST. F - NOT COMPLETELY KNOWN

SAMPLE: $X_1, X_2, \ldots X_n \sim F \in \mathcal{F}$ $\mathcal{F}$ SOME FAMILY OF DISTRIBUTIONS. F BELONGS TO

QUESTION ABOUT F: MEAN, VARIANCE, PROB?

FOR THIS COURSE, $\mathcal{F}$ (F BELONGS TO) IS A PARAMETRIC FAMILY. i.e. $\mathcal{F} = \{ f(x|\theta) ; \theta \in \Theta \}$ θ = PARAMETER Θ = PARAMETER SPACE θ AN UNKNOWN QUANTITY.

QUESTION ABOUT F TRANSLATES TO A QUESTION ON $g(\theta)$, A FUNCTION OF θ.

EX: $\mathcal{F} = \{ f(x|\mu) = N(\mu, \sigma^2)$ DIST., σ^2 KNOWN $\}$ WHAT IS THE VALUE OF μ?

NON-PARAMETRIC SITUATION: FAMILY $\mathcal{F}$ MAY NOT BE INDEXED BY A PARAMETER. EX: $\mathcal{F} = \{$ ALL CONT. DENSITIES $\}$, $\mathcal{F} = \{$ ALL SYMMETRIC DIST. $\}$.

A VERY IMPORTANT ELEMENT OF INFERENCE.

RANDOM SAMPLES: SECS 5.1 & 5.2

$X_1, X_2, \ldots X_n$ OF RVS INDEPENDENT AND IDENTICALLY DISTRIBUTED (iid) EACH FOLLOWING A DENSITY $f(x)$ OR $f(x|\theta)$.

KEY ISSUE OF RANDOM SAMPLING:

IF $f(x) \equiv f(x|\theta)$ THEN THE JOINT PROB. DIST OF $X_1, X_2, \ldots X_n$ IS

$$f(x_1, x_2, \ldots x_n | \theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

FOR EX:

IF $f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$

AND $X_1, X_2, \ldots X_n$ ARE A R.S. WHERE $X_i \sim N(\mu,\sigma^2)$

$\Rightarrow$

$$f(x_1, x_2, \ldots x_n|\mu,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right).$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right). \quad (*)$$

OBS. WHEN THE X'S ARE OBSERVED. $f(x_1 \ldots x_n|\mu,\sigma^2)$ DEFINES THE LIKELIHOOD FUNCTION ON (μ,σ^2) CHAP. 6.

RANDOM SAMPLING REFERS TO SAMPLING FROM INFINITE POPULATIONS*

ANOTHER ISSUE:

SUMMARIES OF RANDOM SAMPLES ARE EXPRESSED BY STATISTICS.

DEF: A STATISTIC IS A FUNCTION $Y = T(X_1 \ldots X_n)$ REAL- OR VECTOR VALUED WHOSE DOMAIN IS THE SAMPLE SPACE OF $(X_1, X_2 \ldots X_n)$.

SAMPLE SPACE = SET OF ALL POSSIBLE VALUES THAT $(X_1, X_2 \ldots X_n)$ MAY TAKE.

EXS:

$T_1(X_1, X_2 \ldots X_n) = \sum_{i=1}^{n} X_i/n = \bar{X}$; $T_3(X_1 \ldots X_n) = X_{(1)}$ $= \min\{X_1, X_n\}$

$T_2(X_1, X_2 \ldots X_n) = \sum_{i=1}^{n} (X_i-\bar{X})^2/n$. $T_4(X_1 \ldots X_n) = X_{(n)}$ $\max\{X_1, \ldots X_n\}$

THE ONLY RESTRICTION ON T, IS THAT IT SHOULD NOT DEPEND ON θ. EX: $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ IS NOT A STATISTIC.

AS A FUNCTION OF $X_1 . X_n$, $T(X_1 . X_2 . X_n)$ IS A R.V.
THE PROPERTIES OF T WILL DEPEND ON THE "PARENTAL" DIST $f(x)$. (REMEMBER TRANSF. OF RVs).

EX3: (THEO 5.2.6). IF $X_1 . X_2 . X_n$ IS A RANDOM SAMPLE WHERE $E(X) = \mu$ AND $VAR(X) = \sigma^2$ THEN
$E(\bar{X}) = \mu$, $VAR(\bar{X}) = \sigma^2/n$, $E(S^2) = \sigma^2$ WHERE $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1)$

($\bar{X}$ AND S^2 ARE UNBIASED TO ESTIMATE μ AND σ^2 RESPECTIVELY)

EX2: (THEO 5.2.7) IF $X_1 . X_2 . X_n$ IS A RANDOM SAMPLE WITH mgf $M_X(t) = E(e^{tX}) \Rightarrow$

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

VERY HELPFUL TO DERIVE THE DIST. OF $\bar{X}$ EASILY

IF $X \sim N(\mu, \sigma^2) \Rightarrow M_X(t) = \exp(\mu t + \frac{\sigma^2 t^2}{2})$

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

$$= \exp\left(\mu t + \frac{\sigma^2}{n}\frac{(t^2)}{2}\right)$$

$\Rightarrow \bar{X} \sim N(\mu, \sigma^2/n)$.

IF $X \sim Exp(\lambda) = f(x:\lambda) = \frac{1}{\lambda} e^{-\lambda^{-1} x}$; $x > 0$ $M_X(t) = \frac{1}{1-\lambda t}$

$\Rightarrow M_{\bar{X}}(t) = \left(\frac{1}{1-\lambda(t/n)}\right)^n$ = GAMMA DIST WITH PARAM. $\alpha = n$
$\beta = \lambda/n$.

(GAMMA $(\alpha.\beta) = f(x|\alpha\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$).

WHAT IF THE mgf DOES NOT EXIST ?

WITH iid $f_{\bar{X}}(x) = n f_{X_1+X_2+ . X_n}(nx)$ (SEE EX 5.5).

ORDER STATISTICS

GIVEN $X_1, X_2 \ldots X_n$ A RANDOM SAMPLES PLACE VALUES IN ASCENDING ORDER

$X_{(1)}, X_{(2)} \ldots X_{(n)}$

$X_{(1)} = \min_{1 \le i \le n} X_i$; $X_{(2)}$ = SECOND SMALLEST $X_i \ldots X_{(n)} = \max_{1 \le i \le n} X_i$

SAMPLE RANGE: $R = X_{(n)} - X_{(1)}$

SAMPLE MEDIAN: M

$X_{(1)}, X_{(2)}, X_{(3)}$ — $M = X_{(2)}$

$X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$ $M = \frac{X_{(2)} + X_{(3)}}{2}$

IN GENERAL

$$M = \begin{cases} X_{((n+1)/2)} & \text{IF } n \text{ IS ODD} \\ \frac{X_{(n/2)} + X_{(n/2+1)}}{2} & \text{IF } n \text{ IS EVEN} \end{cases}$$

M IS A MEASURE OF LOCATION. LESS SENSITIVE THAN THE MEAN TO EXTREME OBSERVATIONS.

MAIN RESULT (THEOREM 5.4.4)

IF $X_{(1)}, X_{(2)} \ldots X_{(n)}$ DENOTE THE ORDER STATISTIC OF A RANDOM SAMPLE $X_1, X_2 \ldots X_n$ OF CONTINUOUS VARIABLES WITH CDF $F_X(x)$ AND pdf $f_X(x)$. THEN, THE pdf OF $X_{(j)}$ IS

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

$\{X_{(j)} \le x\} = \{$AT LEAST j X's ARE ~~[illegible]~~ LESS OR EQUAL TO $x\}$

LET Y = NO. OF X's $\le x$ $\{X_i \le x\}$ SUCCESS EVENT

$\Rightarrow \{X_{(j)} \le x\} = \{Y \ge j\}$ AND $Y \sim \mathrm{BIN}(n, F_X(x))$

$$P[X_{(j)} \le x] = \sum_{k=j}^{n} \binom{n}{k} [F_X(x)]^k [1 - F_X(x)]^{n-k}, \quad k = 0, 1, \ldots n$$

$$f_{X_{(j)}}(x) = \frac{d}{dx} P[X_{(j)} \le x]$$

$$= \sum_{k=j}^{n} \Big[\binom{n}{k} k [F_X(x)]^{k-1} [1-F_X(x)]^{n-k} f_X(x) - \binom{n}{k} [F_X(x)]^{k} (n-k)$$

$$[1-F_X(x)]^{n-k-1} f_X(x) \Big]$$

$$= \underbrace{\binom{n}{j} j [F_X(x)]^{j-1} [1-F_X(x)]^{n-j} f_X(x)}_{\text{WHAT WE WANT}} + \underbrace{\sum_{k=j+1}^{n} \binom{n}{k} k [F_X(x)]^{k-1} [1-F_X(x)]^{n-k} f_X(x)}_{①}$$

$$- \underbrace{\sum_{k=j}^{n-1} \binom{n}{k} (n-k) [F_X(x)]^{k} [1-F_X(x)]^{n-k-1} f_X(x)}_{②}$$

NOTE THAT $\binom{n}{j} j = \frac{n!}{(n-j)!\, j!} j = \frac{n!}{(n-j)!\,(j-1)!}$

CHANGE OF VARIABLE $k = j+1, \ldots n$

$k' = k-1$ so $k' = j, \ldots n-1$

so

$$① = \sum_{k'=j}^{n-1} \binom{n}{k'+1} (k'+1) [F_X(x)]^{k'} [1-F_X(x)]^{n-k'-1} f_X(x)$$

SINCE k' IS A DUMMY VARIABLE MAY REDEFINE k FOR k'

JUST NEED TO CHECK THAT

$$\binom{n}{k} (n-k) = \binom{n}{k+1} (k+1)$$

$$\frac{n!}{k!\,(n-k)!} (n-k) = \frac{n!}{k!\,(n-k-1)!} \quad \text{so} \quad ① - ② = 0$$

PARTICULAR CASES

$j = 1$, $j = n$

$$f_{X_{(1)}}(x) = \frac{n!}{(n-1)!} f_X(x) [1-F_X(x)]^{n-1} = n f_X(x) [1-F_X(x)]^{n-1}$$

$$f_{X_{(n)}}(x) = \frac{n!}{(n-1)!} f_X(x) [F_X(x)]^{n-1} [1-F_X(x)]^{0} = n f_X(x) [F_X(x)]^{n-1}$$

EXAMPLE $X_1, X_2 .\ X_n$ ARE A R.S. FROM $U(0,1)$

$f_X(x) = 1 \quad 0 < x < 1 \qquad f_{X_{(1)}}(x) = n(1-x)^{n-1}$

$F_X(x) = x \qquad f_{X_{(n)}}(x) = n x^{n-1}$

⑥

FOR ARBITRARY j

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

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

WHAT TYPE OF DIST IS THIS? BETA $(j, n-j+1)$

RECALL THAT IF $Z \sim \text{BETA}(\alpha, \beta)$

$$E(Z) = \alpha/(\alpha+\beta) \qquad VAR(Z) = \frac{\alpha \cdot \beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

THEREFORE

$$E(X_{(j)}) = \left(\frac{j}{n+1}\right); \quad VAR(X_{(j)}) = \frac{(j)(n-j+1)}{(n+1)^2(n+2)}.$$

THEOREM 5.4.6

GIVES AN EXPRESSION FOR $f_{X_{(i)}, X_{(j)}}(u,v) \qquad 1 \le i \le j \le n$

IF WE PICK $i=1,\ j=n$ EQ 5.4.7 GIVES

$$f_{X_{(1)}, X_{(n)}}(u,v) = \frac{n!}{(n-2)!} f_X(u) f_X(v) \left[F_X(v) - F_X(u)\right]^{n-2}$$

$$= n.(n-1) f_X(u) f_X(v) \left[F_X(v) - F_X(u)\right]^{n-2}$$

IF FOR EXAMPLE $X_1, X_2 \ldots X_n$ CONFORMS A SAMPLE FROM A UNIFORM $(0,a)$ DIST

$$f_{X_{(1)}, X_{(n)}}(u,v) = n(n-1)\left(\frac{1}{a}\right)\left(\frac{1}{a}\right)\left[\frac{v}{a} - \frac{u}{a}\right]^{n-2}$$

$$= \frac{n(n-1)\,(v-u)^{n-2}}{a^n} \qquad 0 < u < v < a$$

ALSO $f_{X_{(n)}}(v) = n\left(\frac{1}{a}\right)\left(\frac{v}{a}\right)^{n-1} = \frac{n v^{n-1}}{a^n}$

THEREFORE $f_{X_{(1)}|X_{(n)}}(u|v) = \dfrac{f_{X_{(1)}, X_{(n)}}(u,v)}{f_{X_{(n)}}(v)} = \dfrac{n(n-1)(v-u)^{n-2}}{n\, v^{n-1}}$

COND. DIST OF $X_{(1)}$ GIVEN $X_{(n)}$

HW PROB. EX 5.27.

CONVERGENCE CONCEPTS.

WHY? LONG-RUN BEHAVIOR ($n \to \infty$) OF SAMPLE QUANTITIES SUCH AS $\bar{X}$ OR S^2. RELATES TO ASYMPTOTICS. PERMITS APPROXIMATIONS TO "LARGE" n (BUT FIXED) SITUATIONS.

3 TYPES OF CONVERGENCE: i) CONVERGENCE IN PROBABILITY ii) ALMOST SURE CONVERGENCE iii) CONVERGENCE IN DISTRIBUTION

DEF 1: A SEQUENCE OF RANDOM VARIABLES $X_1, X_2, \ldots$ CONVERGES IN PROBABILITY TO X, IF FOR EVERY $\epsilon > 0$

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0 \quad \text{OR} \quad \lim_{n\to\infty} P(|X_n - X| < \epsilon) = 1$$

WE DENOTE THIS BY $X_n \xrightarrow{P} X$

EX: SUP. $X_n \sim N(0, 1/n)$ STA: $X_n \xrightarrow{P} 0$

WHY?

CONSIDER $P(|X_n| < \epsilon) = P(-\epsilon < X_n < \epsilon)$

$$= P\left(-\epsilon/\sqrt{1/n} < Z < \epsilon/\sqrt{1/n}\right) \quad (1) \quad \text{WHERE } Z \sim N(0,1)$$

AS $n \to \infty$

$(1) = P(-\infty < Z < \infty) = 1$

EX 2 (WEAK LAW OF LARGE NUMBERS). LET $X_1, X_2 \ldots X_n$ BE iid RVS WITH $E(X_i) = \mu$ AND $Var(X_i) = \sigma^2$, IF $\bar{X}_n = 1/n \sum_1^n X_i$ THEN, FOR EVERY $\epsilon > 0$, $\lim_{n\to\infty} P(|\bar{X}_n - \mu| < \epsilon) = 1$

<u>PROOF</u> (VIA CHEBYCHEV'S INEQ) FOR EVERY $\epsilon > 0$

$$P(|\bar{X}_n-\mu| \geq \epsilon) = P((\bar{X}_n-\mu)^2 \geq \epsilon^2) \leq \frac{E(\bar{X}_n-\mu)^2}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

$$\Rightarrow P(|\bar{X}_n-\mu| \leq \epsilon) \geq 1-\frac{\sigma^2}{n\epsilon^2}$$ AS $n \to \infty$ PROB $\to 1$.

FOR S^2: $S^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2$ CAN WE PROVE A WLLN?

$$P(|S_n^2-\sigma^2| \geq \epsilon) \leq \frac{E(S_n^2-\sigma^2)^2}{\epsilon^2} = \frac{VAR(S_n^2)}{\epsilon^2} \to 0$$ AS LONG AS $VAR(S_n^2) \to 0$.

ALMOST SURE CONVERGENCE:

DEF 2.: A SEQUENCE OF RANDOM VARIABLES $X_1, X_2, \ldots$ CONVERGES ~~IN~~ ALMOST SURELY TO X, IF EVERY $\epsilon > 0$

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

WE DENOTE THIS BY $X_n \xrightarrow{a.s.} X$

RESULT: ALMOST SURE CONVERGENCE IMPLIES CONVERGENCE IN PROBABILITY. BUT NOT THE OPPOSITE

EX: 5.5.8

CONSIDER A SAMPLE SPACE $S = [0,1]$ WITH A $U(0,1)$ DISTRIBUTION.

FOR ANY POINT $s \in S$, WE DEFINE

$X_1(s) = s + I_{[0,1]}(s)$; $X_2(s) = s + I_{[0,1/2]}(s)$,

$X_3(s) = s + I_{[1/2,1]}(s)$; $X_4(s) = s + I_{[0,1/3]}(s)$,

$X_5(s) = s + I_{[1/3,2/3]}(s)$; $X_6(s) = s + I_{[2/3,1]}(s)$;

$X_7(s) = s + I_{[0,1/4]}(s)$; $X_8(s) = s + I_{[1/4,1/2]}(s)$;

LETS MAKE $X(s) = s$ (IDENTITY FUNCTION).

$P(|X_n(s) - X(s)| \geq \epsilon) = P(I_n) = \text{LENGTH}(I_n)$

BUT AS $n \to \infty$ length $(I_n) \to 0 \Rightarrow X_n \xrightarrow{P} X$.

FOR A.S., WE NEED $X_n(s) \to s$ EXCEPT ON A SET WITH MEASURE ZERO

NOTICE THAT FOR ANY s, $X_n(s)$ IS $s+1$, OR s

TAKE $s = 1/2$ $X_1(1/2) = 3/2$, $X_2(1/2) = 3/2$, $X_3(1/2) = 3/2$

$X_4(s) = 1/2$, $X_5(s) = 3/2$.

NOTICE THAT THE SUBSEQUENCE $X_{n_i}(s) = s + I_{(0,1/n)}(s)$ CONVERGES BOTH IN PROB. AND A.S.

FOR LARGE ENOUGH n, $X_n(s) = s$

JAN 29, 2002 LAST. CONV IN P, CONV. a.s. ANOTHER EX.

STRONG LAW OF LARGE NUMBERS. LET $X_1, X_2, \ldots$ BE iid RVs WITH MEAN μ AND VARIANCE σ^2. THEN, FOR EVERY

$\epsilon > 0$ $\quad P(\lim_{n\to\infty} |\bar{X}_n - \mu| < \epsilon) = 1$ $\quad (\bar{X}_n \xrightarrow{a.s.} \mu)$

LAST FORM OF CONVERGENCE.

CONVERGENCE IN DIST.

DEF 3. A SEQUENCE OF RANDOM VARIABLES $X_1, X_2, \ldots$ CONVERGES IN DIST. TO A RANDOM VARIABLE X IF $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ DENOTED BY $X_n \xrightarrow{D} X$

EX: $Z \sim N(0,1)$ $\quad X_1 = Z,\ X_2 = -Z,\ X_3 = Z \ldots$

NOTICE THAT $-Z \sim N(0,1)$ ~~BUT~~ AND THEN $X_n \xrightarrow{D} Z$

CAN THIS SEQUENCE CONVERGE IN PROBABILITY TO Z?

NOTICE THAT: $P(|X_n - Z| \geq \epsilon) = \begin{cases} 0 & \text{IF } X_n = Z \ (n \text{ ODD}) \\ P[2|Z| \geq \epsilon] & \text{IF } X_n = -Z \ (n \text{ EVEN}). \end{cases}$

$\Rightarrow$ X_n DOES NOT CONVERGE IN PROBABILITY.

<u>TEO</u> IF $X_1, X_2, \ldots X_n$ CONVERGES IN PROB. TO X, THE SEQUENCE ALSO CONVERGES IN DIST. TO X.

PROOF. (EX. 5.40) CONT. CASE. WE CAN SHOW THAT

$P(X \leq t-\epsilon) \leq P(X_n \leq t) + P(|X_n - X| \geq \epsilon)$ (1)

SINCE $X_n \leq t$ & $|X_n - X| \geq \epsilon$ ~~AND~~ $X - X_n \leq -\epsilon + X_n \leq t$

$\Rightarrow X \leq t - \epsilon$

IN THE SAME MANNER.

$P(X_n \leq t) \leq P(X \leq t+\epsilon) + P(|X_n - X| \geq \epsilon)$ (2)

COMBINING (1) & (2), WE HAVE

$P(X \leq t-\epsilon) - P(|X_n - X| \geq \epsilon) \leq P(X_n \leq t) \leq P(X \leq t+\epsilon) + P(|X_n - X| \geq \epsilon)$.

AS $n \to \infty$ $\quad P(X \leq t-\epsilon) \leq P(X_n \leq t) \leq P(X \leq t+\epsilon)$

$$F_X(t-\epsilon) \leq F_{X_n}(t) \leq F_X(t+\epsilon)$$

$\Rightarrow \lim F_{X_n}(t) = F_X(t)$

FAMOUS EXAMPLE ~~ON~~ CONVERGENCE IN DIST.

CENTRAL LIMIT THEOREM: IF $X_1, X_2, \ldots$ ARE iid RVS. WITH $E(X_i)=\mu$ AND $Var(X_i)=\sigma^2$. IF $Z_n=\sqrt{n}(\bar{X}_n-\mu)/\sigma$ AND $F_{Z_n}(z)$ IS THE DIST. FUNCTION OF Z_n THEN

$$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z) \qquad \text{WITH}$$

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2}\,dx$$

BOOK OFFERS PROOF WHEN $M_{X_i}(t)$ IN A NEIGHBORHOOD OF 0. BUT THE RESULT IS MORE GENERAL

EX. (TYPICAL OF 345)

Suppose $X_1, X_2, \ldots X_{30}$ ARE iid $X_i \sim \text{Poisson}(\lambda)$

$f(x_i|\lambda) = \dfrac{\lambda^{x_i} e^{-\lambda}}{x_i!}$; $i=1,2,3\ldots n$. Suppose $\lambda = 10$

$P(\sum_{i=1}^{n} X_i < 325)$?

By MGF $\quad M_{\sum X_i}(t) = E(e^{t\sum X_i}) = [M_X(t)]^n$

If $X \sim \text{Poisson}(\lambda) \quad M_X(t) = e^{\lambda(e^t-1)} \Rightarrow [M_X(t)]^n = e^{n\lambda(e^t-1)}$

$\sum X_i \sim \text{Poisson}(n\lambda) \qquad n\lambda = 300$

EXACT CALCULATION $\quad P(\sum_{i=1}^{30} X_i < 325) = 0.9280727$

$$P(\sum_{1}^{30} X_i < 325) = P(\bar{x} < 10.83333) \approx P\left(Z < \frac{.83333}{\sqrt{10/30}}\right) =$$

0.9255419

ONE PROBLEM WITH CLT IS THAT WE REALLY DON'T KNOW HOW LARGE NEEDS n TO BE FOR THE APPROX. TO WORK. (DEPENDS OF $f(x)$).

SUPPOSE NOW THAT WE CARE ABOUT $g(\bar{x})$, SOME FUNCTION OF $\bar{x}$? DO WE HAVE SOME SORT OF CLT FOR $g(\bar{x})$?

IN THE POISSON EXAMPLE, IT SEEMS NATURAL TO ESTIMATE λ (MEAN) WITH $\bar{X}$. BUT NOW IF WANT TO ESTIMATE $e^{-\lambda} = P(X=0)$ AN ESTIMATOR WOULD BE $e^{-\bar{X}} = g(\bar{x})$.

LIMIT DIST. FOR $e^{-\bar{X}}$?

<u>THE DELTA METHOD</u>.

SUPPOSE ~~WE~~ WE HAVE A R.V. (STATISTIC) T SUCH THAT $E(T) = \theta$ AND $g'(\,)$ EXISTS AND IS DIFFERENT OF ZERO

t AN OBSERVED VALUE OF T.

LETS TAKE THE TAYLOR SERIES EXP OF g AROUND θ.

$$g(t) = g(\theta) + g'(\theta)(t-\theta) + \text{REMINDER}.$$

$$\Rightarrow \quad g(t) \approx g(\theta) + g'(\theta)(t-\theta)$$

IF WE TAKE EXPECTATIONS,

$$E(g(T)) \approx g(\theta) + g'(\theta)\,E(T-\theta) = g(\theta)$$

SIMILARLY

$$E(g(T) - g(\theta))^2 \approx (g'(\theta))^2\,E(T-\theta)^2$$

$$\Rightarrow \quad \mathrm{VAR}(g(T)) \approx g'(\theta)^2\,\mathrm{VAR}(T)$$

WE OBTAINED APPROX. EXPRESSIONS FOR $E(g(T))$ AND $\mathrm{VAR}(g(T))$

IN THE EX, $\theta = \lambda$

$$E(e^{-\bar{X}}) \approx e^{-\lambda} \text{ AND } \mathrm{VAR}(e^{-\bar{X}}) \approx (e^{-2\lambda})\left(\frac{\lambda}{n}\right)$$

$g(T)$ IS "APROX." UNBIASED

IN FACT, IF T_n IS A SEQUENCE OF RVs WHERE $\sqrt{n}\,(T_n - \theta) \xrightarrow{d} N(0, \sigma^2)$. IF g IS A FUNCTION WHERE g' EXISTS AND IS NOT 0 THEN:

$$\sqrt{n}\,(g(T_n)-g(\theta)) \xrightarrow{d} N(0, \sigma^2(g'(\theta))^2)$$

IN TERMS OF THE EXAMPLE

$$\sqrt{n}\,(e^{-\bar{X}}-e^{-\lambda}) \xrightarrow{d} N\left(0, e^{-2\lambda}\left(\frac{\lambda}{n}\right)\right).$$

USUALLY σ^2 AND θ ARE UNKNOWN, HENCE $\sigma^2 g'(\theta)^2$ IS UNKNOWN

IF WE CAN FIND V_n SUCH THAT $V_n \xrightarrow{P} \sigma^2 g'(\theta)^2$

THEN SLUTSKY'S THEOREM (PAG. 240) GUARANTEES THAT:

$$\underbrace{\frac{\sqrt{n}\,(g(T_n)-g(\theta))}{\sqrt{V_n}}}_{\text{DOES NOT INVOLVE } \theta \text{ IN THE DENOM.}} \xrightarrow{d} N(0,1).$$

IN THE EXAMPLE POISSON, BY THE WLLN $\bar{X} \xrightarrow{P} \lambda$

BY HW. EX. PROB. $V_n = e^{-2\bar{X}}\left(\frac{\bar{X}}{\lambda}\right) \xrightarrow{P} e^{-2\lambda}\left(\frac{\lambda}{n}\right)$

THEN

$$\sqrt{n}\,(e^{-\bar{X}}-e^{-\lambda}) \Big/ \sqrt{e^{-2\bar{X}}\left(\frac{\bar{X}}{\lambda}\right)} \xrightarrow{D} N(0,1).$$

JUSTIFIES THE USE OF A "PLUG IN" ESTIMATOR FOR A VARIANCE

FINALLY, WHAT IF $g'(\theta)=0$

SECOND ORDER EXP. ON $g(T_n)$

$$g(T_n) = g(\theta) + g'(\theta)(T_n-\theta) + \frac{1}{2}g''(\theta)(T_n-\theta)^2 + R.$$

$$\Rightarrow g(T_n) \simeq \frac{1}{2}g''(\theta)(T_n-\theta)^2 + g(\theta)$$

$$\Rightarrow n(g(T_n)-g(\theta)) \approx n\, g''(\theta)(T_n-\theta)^2/2$$

SO IF $\sqrt{n}(T_n-\theta) \xrightarrow{d} N(0,\sigma^2) \Rightarrow n(T_n-\theta)^2 \overset{d}{\sim} \sigma^2\chi^2_{(1)}$

$$n\,(g(T_n)-g(\theta)) \xrightarrow{d} \sigma^2\,\frac{g''(\theta)}{2}\,\chi^2_{(1)}$$

GENERATING RANDOM SAMPLES.

IN INFERENCE, WE ARE INTERESTED IN THE DISTRIBUTION OF $T(X_1, X_2, \ldots X_n)$ A "STATISTIC"

SUPPOSE IT IS VERY HARD TO DEAL WITH THE DIST OF T WITH mgf OR CONVERGENCE CONCEPTS. (FOR EX. AN ORDER STATISTIC).

IDEA:

SINCE $X_1, X_2 \ldots X_n$ ARE iid WHERE $X_i \sim f(x|\theta)$

WE CAN "GENERATE" WITH THE HELP OF THE COMPUTER VALUES FOR $X_1, X_2, \ldots X_n$. TO OBTAIN VALUES FOR $T(X_1, \ldots X_n)$

EX:

LET $X_1, X_2, \ldots X_{20}$ BE iid $N(0,1)$ RVs.

AND $T = X_{(n)} = \max\{X_1 \ldots X_{20}\}$

HARD TO DEAL WITH THE DIST OF T BECAUSE

$$F_T(t) = P(X_{(n)} \leq t) = \prod_{i=1}^{n} P(X_i \leq t) = [\Phi(t)]^n$$

AND $\Phi(t)$ DOES NOT HAVE A CLOSED FORM. EXPRESSION.

GENERATE VALUES.

$1.-\ X_1^{(1)}, X_2^{(1)}, \ldots X_{20}^{(1)} \longrightarrow X_{(20)}^{(1)}$ AND REPEAT AND OVER

$2.-\ X_1^{(2)}, X_2^{(2)}, \ldots X_{20}^{(2)} \longrightarrow X_{(20)}^{(2)}$

$\vdots$

$1000.-\ X_1^{(1000)}\ X_2^{(1000)} \ldots X_{20}^{(1000)} \longrightarrow X_{(20)}^{(1000)}$

SUMMARIZE $X_{(20)}^{(1)}, X_{(20)}^{(2)}, \ldots X_{(20)}^{(1000)}$.

LOOK AT HANDOUT. FIG (a)

IN FACT, IF WE WANT TO APPROXIMATE $E(X_{(20)})$ AND $E(X_{(20)}^2)$

$$E(X_{(20)}) \approx \frac{1}{1000} \sum_{i=1}^{1000} X_{(20)}^{(i)} = 1.867 \quad \text{(WITH MY SIMULATION)}$$

$$E(X_{(20)}^2) \approx \frac{1}{1000} \sum_{i=1}^{1000} (X_{(20)}^{(i)})^2 = 3.786$$

ISSUE: HOW DO WE GENERATE X WITH A pdf $f(x|\theta)$. ?

START WITH AN ALGORITHM THAT GENERATES U, U FOLLOWS A $U(0,1)$ DISTRIBUTION.

FOR X CONTINUOUS, WE GENERATE X BY

$$X = F^{-1}(U)$$

WHERE F^{-1} IS THE ~~[illegible]~~ p. INVERSE CDF OF F

PROB. INTEGRAL TRANSFORM.

EX. IF X IS $Exp(\lambda)$ THEN $f(x|\lambda) = \frac{1}{\lambda} e^{-x/\lambda}$; $x > 0$

SO $F(x|\lambda) = 1 - e^{-x/\lambda}$; IF WE MAKE

$U = 1 - e^{-x/\lambda} \Rightarrow e^{-x/\lambda} = 1 - U$

$\Rightarrow -\frac{x}{\lambda} = \log(1-U) \Rightarrow X = -\lambda \log(1-U)$

PROB. F^{-1} MAY NOT HAVE A CLOSED FORM. (BOX-MÜLLER).

FOR X DISCRETE:

SUP X TAKES 3 VALUES: x_1 WITH PROB. p_1, x_2 WITH PROB. p_2 AND x_3 WITH PROB. p_3.

TO GENERATE ONE VALUE OF X:

i) GENERATE U FROM A $U(0,1)$

ii) IF $U \leq p_1 \longrightarrow X = x_1$

IF $\underbrace{p_1}_{F(x_1)} < U \leq \underbrace{p_1+p_2}_{F(x_2)} \longrightarrow X = x_2$.

OR IF $P_1+P_2 \leq U \leq P_1+P_2+P_3 \rightarrow X = X_3$.

MOST STATISTICAL PACKAGES HAVE A SIMULATOR FOR WELL ESTABLISHED $f(x|\theta)$. (SEE FIGS. (b)-(d)).

STILL, IF $f(x|\theta)$ IS NOT AVAILABLE, WE MAY USE INDIRECT METHODS.

$X \sim DEXP(\mu=0)$. SO $f(x|\mu) = \frac{1}{2} e^{-|x-\mu|} \quad -\infty < x < \infty$

$(\sigma=1)$ SUPPOSE WE WANT TO GENERATE VALUES OF X.

PROPOSED AUXILIARY DENSITY. (CANDIDATE)

$$f(y|\mu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y-\mu)^2\right) \quad -\infty < y < \infty$$

GENERATE $Y \sim N(0,1)$, SET $X_0 = Y$

FOR $i = 1, 2, \ldots$

1.- GENERATE $U_i \sim U(0,1)$; $Y_i \sim N(0,1)$.

$$\rho_i = \min\left\{1, \frac{\exp(-\frac{1}{2}|y_i|)\exp(-\frac{1}{2}(x_i)^2)}{\exp(-\frac{1}{2}y_i^2)\exp(-\frac{1}{2}|x_i|)}\right\}$$

2.- SET

$$x_i = \begin{cases} y_i & \text{IF } U_i \leq \rho_i \\ x_{i-1} & \text{IF } U_i > \rho_i \end{cases}$$

THE KEY THING IS THAT AS $i \rightarrow \infty$ $X_i \rightarrow X$

THIS LEADS TO THE METROPOLIS-HASTINGS ALGORITHM.