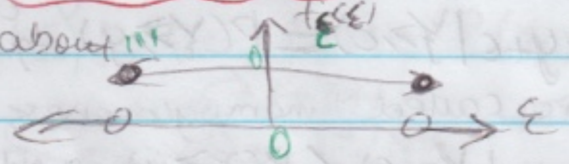
Lecture 12

Laplace First distributed in 1774 calling it "First law of errors", His ~~Firs~~ context was measurement, when you measure a quantity v you measure it with error, epsilon ε so that your measurement is:

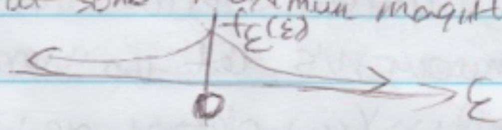
$M = v + (\text{epsilon})\,\varepsilon$

What makes a good distribution for error, epsilon ε? The expectation should be zero (0), and should be symmetric.

How about !!!



This is not very good, It should have property that probability of error should decrease with its magnitude. Also, why should it stop at some maximum magnitude



Another good property is that density should be decreasing in magnitude of error:

Laplace assumed for all positive errors that $f''_{\mathcal{E}}(\varepsilon) = f'_{\mathcal{E}}(\varepsilon)$ (ε small, $\mathcal{E}$ capital)

$\Rightarrow f(\varepsilon) = ce^{-d\varepsilon} \Rightarrow \mathcal{E} \sim \text{Laplace}(0,1)$

$X \sim \text{Exp}(1) = e^{-x}\mathbb{1}_{x \geq 0}$ and $Y = g(X) = \frac{1}{\lambda}X^{\frac{1}{k}}$ s.t. $k, \lambda > 0$ (scale) $\sim f_Y(y) = ?$

First step: get inverse function, second step: abs (absolute) inverse derivative.

$y = \frac{1}{\lambda}x^{\frac{1}{k}} \Rightarrow \lambda y = x^{\frac{1}{k}} \Rightarrow x = (\lambda y)^k = \lambda^k y^k = g^{-1}(y)$ all positive

Second step / Test Question: $\left|\frac{d}{dy}[g^{-1}(y)]\right| = \left|\frac{d}{dy}[\lambda^k y^k]\right| = |k\lambda^k y^{k-1}| = |k\lambda^k||y^{k-1}| = k\lambda^k y^{k-1}$

$\Rightarrow f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}[g^{-1}(y)]\right| = e^{-(\lambda y)^k}\mathbb{1}_{\lambda^k y^k \geq 0} \cdot k\lambda^k y^{k-1}$

$= k\lambda^k y^{k-1} e^{-(\lambda y)^k}\mathbb{1}_{y \geq 0} = k\lambda(\lambda y)^{k-1} e^{-(\lambda y)^k}\mathbb{1}_{y\geq 0}$ ($y^k \geq 0$, $y \geq 0$) $= \text{Weibull}(k, \lambda)$

1951

17:35

Weibull is very famous waiting time/survival rv model and it's used eg. in insurance companies to price life insurance (I think)

$$\text{Weibull}(\overset{k=1}{1}, \lambda) = (1)\lambda(\lambda y)^{1-1} e^{-(\lambda y)^1} \mathbb{1}_{y\geq 0} = \lambda e^{-\lambda y} \mathbb{1}_{y\geq 0} = \text{Exp}(\lambda)$$

L as a generalize of Exponential

K parameter is really "cool" Here's a property of Weibull rv's under different values of K:

$K=1$ $P(Y \geq y+c \mid Y \geq c) = P(Y \geq y)$ eg. $c=14,\ y=3$ $P(Y\geq 17 \mid Y \geq 14) = P(Y \geq 3)$

- this equality is called "memorylessness"

$K>1$ $P(Y \geq y+c \mid Y \geq c) < P(Y \geq y)$ eg. old lifespan of human, waiting For Bus

$K<1$ $P(Y \geq y+c \mid Y \geq c) > P(Y \geq y)$ eg startup company life spam natal lifespam

Order statistics P(160) let $X_1, X_2 \cdots X_n$ be a collection of continuous rv's, let the "order statistics" be the rv's: $X_{(1)}, X_{(2)} \cdots X_{(n)}$ defined as:

$X_{(1)} := \min\{X_1, X_2 \cdots X_n\}$

$\vdots$

$X_{(k)} := k^{th}$ largest of $X_1 \cdots X_n$

$\vdots$

$X_{(n)} := \max\{X_1, X_2, \ldots, X_n\}$

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

e.g. $X_1 = 9$. $X_2 = 2$. $X_3 = 12$

$X_4 = 7$

$X_{(1)} = X_2$

$X_{(2)} = 7$

$X_{(3)} = 9$

$X_{(4)} = 12$

$r = 12 - 2 = 10$

We want to find Both CDF and PDF of k^{th} order statistic. We will build this up in stages. The First thing we'll do is find the CDF and PDF of maximum.

certain value

$n=5$

CDF: $F_{X_{(n)}}(x) = P(\underbrace{X_{(n)} \le x}_{\text{event}}) = P(X_1 \le x, X_2 \le x, \dots X_n \le x)$

Assume If $X_1 \dots X_n$ independent & iid (All CDF same)

PDF only for iid cases $\overset{\text{if iid}}{=} P(X_1 \le x) P(X_2 \le x) \dots P(X_n \le x) = \prod_{i=1}^{n} F_{X_i}(x) \overset{\text{CDF}}{=} F_X(x)^n$

PDF: $f_{X_n}(x) := \frac{d}{dx}[F_X(x)^n] = n \underbrace{f(x)}_{=F'(x)} F(x)^{n-1}$

next thing we'll do is to find CDF and PDF of minimum

$F_{X_{(1)}}(x) = P(X_{(1)} \le x) = 1 - P(X_{(1)} > x) \overset{\text{minimum}}{=} 1 - P(X_1 > x, X_2 > x, \dots X_n > x)$

$\overset{\text{if independent}}{=} 1 - P(X_1 > x) \dots P(X_n > x) =$

$= 1 - \prod_{i=1}^{n} \underbrace{P(X_i > x)}_{1-F_{X_i}(x)} \overset{\text{CDF}}{=} 1 - \prod_{i=1}^{n} (1 - F_{X_i}(x)) \overset{\text{if iid CDF are same}}{=} 1 - (1 - F(x))^n$

$f_{X_{(1)}}(x) \overset{\text{iid}}{=} \frac{d}{dx}[1 - (1 - \overset{\text{CDF}}{F}(x))^n] = n f(x) (1 - F(x))^{n-1}$

next thing we will do is assume $n=10$ and derive the $k=4$th order static's CDF and PDF Before get there, let's find probability that first four numbers are less than x and last six number are greater than x

$= P(X_1 \le x, \dots X_4 \le x, X_5 > x, \dots X_{10} > x)$

$\overset{\text{if independent}}{=} \prod_{i=1}^{4} \overset{\text{CDF}}{F_{X_i}(x)} \prod_{i=5}^{10} (1 - \overset{\text{survival}}{F_{X_i}(x)}) \overset{\text{if iid}}{=} F(x)^4 (1 - F(x))^6$

Let's find probability 4 of 10 are below x and remaining are above x. let S be subset of size 4 of index set $\{1, 2, \dots 10\}$

$= \sum_{\text{all } S} P(X_{S_1} \le x, \dots, X_{S_4} \le x, X_{S_1^c} > x, \dots X_{S_6^c} > x)$

$\overset{\text{if independent}}{=} \sum_{\text{all } S} \prod_{i=1}^{4} F_{X_{S_i}}(x) \prod_{i=1}^{6} (1 - F_{X_{S_i^c}}(x)) \overset{\text{if iid}}{=} \sum_{\text{all } S} F(x)^4 (1 - F(x))^6$

$$= \binom{10}{4} F(x)^4 (1-F(x))^6$$

Now let's derive CDF for $k=4$th order statistic

$$F_{X_{(4)}}(x) = \underbrace{P(X_{(4)} \le x)}_{\text{event}} = P\left(\begin{array}{c}\text{a subset of } 4\ X_i\text{'s} \le x, \text{ the} \\ \text{Remaining } 6 \text{ are} > x\end{array}\right)$$

$$+ P\left(\begin{array}{c}\text{a shuset of } 5\ X_i\text{'s} \le x \text{ and} \\ \text{Remaing } 5 \text{ are} > x\end{array}\right) +$$

$$\vdots$$

$$+ P(\text{All } 10\ X_i\text{'s} \le x)$$

$$\overset{\text{If iid}}{=\!=} \binom{10}{4} F(x)^4 (1-F(x))^6 + \binom{10}{5} F(x)^5 (1-F(x))^5$$

$$+ \cdots + \binom{10}{10} F(x)^{10} (1-F(x))^{10-10}$$

$$= \sum_{j=4}^{10} \binom{10}{j} F(x)^j (1-F(x))^{10-j}$$

For iid continuous rv's $X_1, \dots X_n$, the CDF and PDF for kth order statistic is:

CDF $F_{X_{(k)}}(x) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j (1-F(x))^{n-j}$

PDF $f_{X_{(k)}}(x) = \frac{d}{dx}\left[\sum_{j=k}^{n} \binom{n}{j} \overbrace{F(x)^j}^{u} \overbrace{(1-F(x))^{n-j}}^{v}\right]$ $\quad \frac{d}{dx}(uv) = uv' + u'v$

$$= \sum_{j=k}^{n} \binom{n}{j} \frac{d}{dx}\left[F(x)^j (1-F(x))^{n-j}\right]$$

$u' = j f(x) F(x)^{j-1} \quad v' = -(n-j) f(x) (1-F(x))^{n-j-1}$