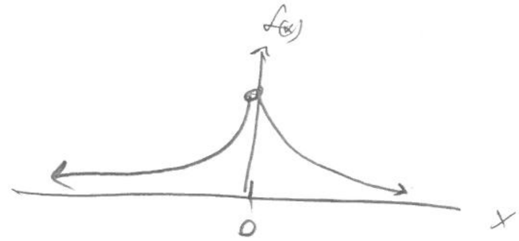
Lec 11 Math 621 10/23/17

$$\text{Laplace}(0,1) = \frac{1}{2} e^{-|x|}$$



Laplace published this in 1774 calling it "first law of errors".

Imagine you try to measure a quantity v but your measuring device has random error, ε. So your measurement M looks like

$M = v + \varepsilon$. So $M \approx v$. What makes a good error distr? It would make sense

for $E[\varepsilon] = 0 \Rightarrow E(M) = v$ i.e. M is "unbiased" 369/638

Similarly, $\text{Med}[\varepsilon] = 0$. So 50% of the time you measure $< v$
and 50% " " " " $> v$.

Also $f(\varepsilon) = f(-\varepsilon)$. So measuring $> \varepsilon$ away is just as probable as measuring $< -\varepsilon$ away

How about?

It would make sense for larger errors to be less probable.

So $f'(\varepsilon) < 0$ if $\varepsilon > 0$ & $f'(\varepsilon) > 0$ if $\varepsilon < 0$.

If $f''(\varepsilon) = f'(\varepsilon) \Rightarrow f(\varepsilon) = ce^{-dx} \Rightarrow \text{Laplace}(0,1)$.

$X \sim \text{Exp}(1) = e^{-x}\mathbb{1}_{x\geq 0}$, $\quad Y = \frac{1}{\lambda} X^{\frac{1}{k}}$ where $\lambda, k > 0$

$$\Rightarrow \lambda Y = X^{\frac{1}{k}} \Rightarrow X = (\lambda Y)^k = \lambda^k Y^k = g^{-1}(Y)$$

$$\left|\frac{d}{dy}\left[g^{-1}(y)\right]\right| = \lambda^k\left(k y^{k-1}\right) = k\lambda^k y^{k-1}$$

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

special case

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

Weibull is another generalization of the exponential. It is also a waiting time / survival distr.

$$F(y) = \int_0^y k\lambda \underbrace{(\lambda t)^{k-1}}_{\lambda^k t^{k-1}} e^{-(\lambda t)^k}\,dt = \int_0^{(\lambda y)^k} \cancel{k\lambda^k t^{k-1}}\, e^{-u}\,\frac{1}{\cancel{\lambda^k k t^{k-1}}}\,du = \left[-e^{-u}\right]_0^{(\lambda y)^k} = 1 - e^{-(\lambda y)^k}$$

let $u = (\lambda t)^k = \lambda^k t^k$ $\quad \frac{du}{dt} = \lambda^k k t^{k-1} \Rightarrow dt = \frac{1}{\lambda^k k t^{k-1}}\,du$ $\qquad \Rightarrow \bar{F}(y) = e^{-(\lambda y)^k}$

$t = y \Rightarrow u = (\lambda y)^k$, $\; t = 0 \Rightarrow u = 0$

Cool property:

$c > 0$

$$P(Y \geq y+c \mid Y \geq c) = \frac{\overbrace{P(Y\geq y+c \;\&\; Y \geq c)}^{P(Y\geq y+c)}}{P(Y\geq c)} = \frac{e^{-(\lambda(y+c))^k}}{e^{-(\lambda c)^k}} = e^{\lambda^k(c^k - (y+c)^k)}$$

If $k = 1$, $\; P(Y \geq y+c \mid Y \geq c) = P(Y \geq y)$ memoryless! Because it's an exponential!

If $k > 1$, $\; P(Y \geq y+c \mid Y \geq c) < P(Y \geq y)$ gets less probable as c increases. Example?

If $k < 1$, $\; P(Y \geq y+c \mid Y \geq c) > P(Y \geq y)$ gets more prob. as c increases. Example?

e.g. if $k=2$

$$e^{\lambda^2(c^2-(y+c)^2)} < e^{-\lambda^2 y^2}$$

$\Rightarrow \cancel{\lambda^2}\left(c^2-(y^2+2yc+c^2)\right) < -\cancel{\lambda^2}y^2$

$\Rightarrow \cancel{c^2}-\cancel{y^2}-2yc-\cancel{c^2} < \cancel{-y^2}$

$\Rightarrow -2yc < 0$ ✓ since $y,c>0$

e.g. if $k=\frac{1}{2}$

$\Rightarrow \sqrt{c}-\sqrt{y+c} > \sqrt{y}$

$\Rightarrow \sqrt{c}-\sqrt{y} > \sqrt{y+c}$

$\Rightarrow \cancel{c+y}+2\sqrt{cy} > \cancel{y+c}$

$\Rightarrow 2\sqrt{cy} > 0$ since $cy>0$ ✓

General proof on HW. Weibull is super flexible model!

Order Statistics (p160)

Let $X_1, X_2, \dots, X_n$ be a collection of cont. r.v.'s then let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be called the "Order Statistics" which are r.v.'s defined as

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

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

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

Also... let

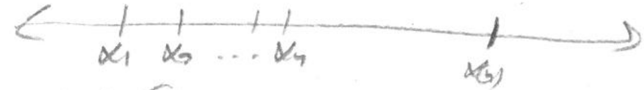
$R := X_{(n)} - X_{(1)}$

"the range"

For realizations $x_1 = 9,\ x_2 = 2,\ x_3 = 12,\ x_4 = 7$

$\Rightarrow x_{(1)} = 2,\ x_{(2)} = 7,\ x_{(3)} = 9,\ x_{(4)} = 12$ & $r = 12-2 = 10$

Let's find the PDF & CDF of the maximum, $X_{(n)}$,



$$F_{X_{(n)}}(x) := P(X_{(n)} \le x) = P(X_1 \le x \;\&\; X_2 \le x \;\&\; \ldots \;\&\; X_n \le x)$$

if $X_1, \ldots, X_n$ ind

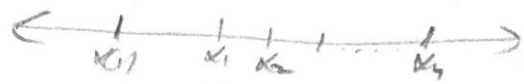
$$= \prod_{i=1}^{n} P(X_i \le x) = \prod F_{X_i}(x)$$

if $X_1, \ldots, X_n$ iid

$$= P(X \le x)^n = F_X(x)^n$$

$$f_{X_{(n)}}(x) = \frac{d}{dx}\left[F_X(x)^n\right] = n f(x) F_X(x)^{n-1} \quad \text{if iid}$$

Let's find the PDF & CDF of the minimum, $X_{(1)}$,



$$F_{X_{(1)}} := P(X_{(1)} \le x) = 1 - P(X_{(1)} > x) = 1 - P(X_1 > x \;\&\; X_2 > x \;\&\; \ldots \;\&\; X_n > x)$$

if $X_1, \ldots, X_n$ ind

$$= 1 - \prod P(X_i > x) = 1 - \prod (1 - F_{X_i}(x))$$

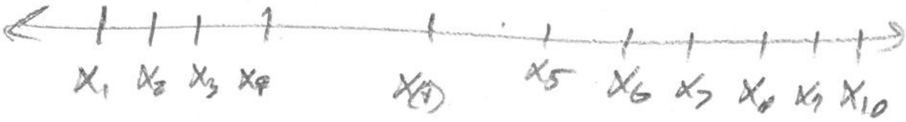
if $X_1, \ldots, X_n$ iid

$$= 1 - P(X > x)^n = 1 - (1 - F(x))^n$$

$$f_{X_{(1)}}(x) = \frac{d}{dx}\left[F_{X_{(1)}}(x)\right] = -(-f(x))\, n\, (1 - F(x))^{n-1} = n f(x) (1 - F(x))^{n-1}$$

Let's get the PDF & CDF of $X_{(4)}$.

Let's consider $n=10$ to get intuition and $X_{(4)}$. Consider:



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

if $X_1, \dots, X_{10}$ ind

$$= \prod_{i=1}^{4} P(X_i \le x) \prod_{i=5}^{10} P(X_i > x) = \prod_{i=1}^{4} F_{X_i}(x) \prod_{i=5}^{10} (1 - F_{X_i}(x))$$

if $X_1, \dots, X_{10}$ iid

$$= F(x)^4 (1 - F(x))^6$$

Consider $P($any 4 X_i's $\le x$ and the other 6 X_i's $> x)$

$$= \sum_{\substack{\text{all subsets} \\ \text{of size 4}}} P(X_{S_1} \le x \dots X_{S_4} \le x \,\&\, X_{S_5} > x \,\&\, \dots \,\&\, X_{S_{10}} > x)$$

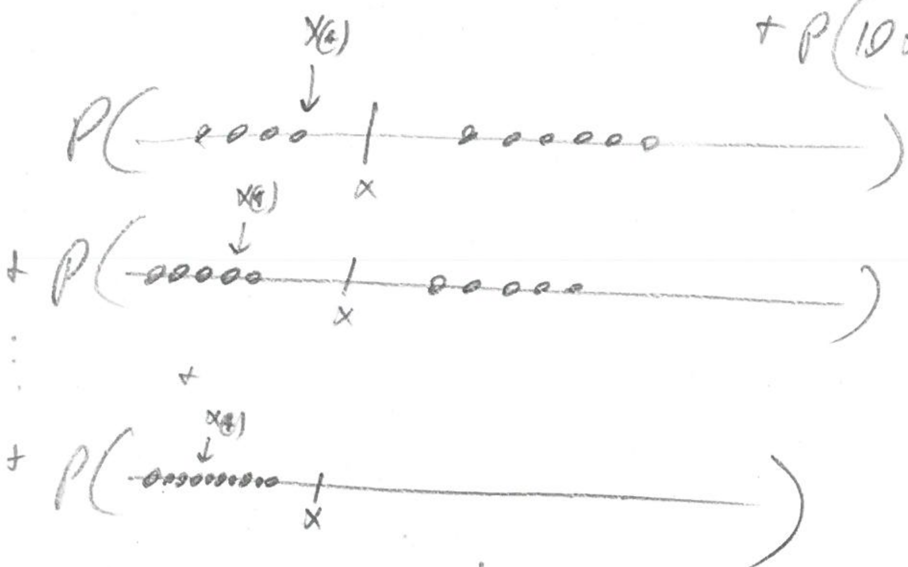
if ind

$$= \sum \prod_{i=1}^{4} F_{X_{S_i}}(x) \prod_{i=5}^{10} (1 - F_{X_{S_i}}(x))$$

if iid

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

$$F_{X_{(4)}}(x) := P(X_{(4)} \le x) = P(4\ x\text{'s are below } x \text{ and } 6\ x\text{'s above})$$
$$+ P(5\ x\text{'s} \ \text{''} \ \text{''} \ \text{and } 5\ x\text{'s above})$$
$$+ P(6\ x\text{'s} \ \ldots\ldots \ 4\ x\text{'s} \ldots)$$
$$\vdots$$
$$+ P(10\ x\text{'s} \ \ldots\ldots \ 0\ x\text{'s} \ldots)$$



x iid

$$= \binom{10}{4} F(x)^4 (1-F(x))^6$$
$$+ \binom{10}{5} F(x)^5 (1-F(x))^5$$
$$\vdots$$
$$+ \binom{10}{10} F(x)^{10} (1-F(x))^0 = \sum_{j=k}^{10} \binom{10}{j} F(x)^j (1-F(x))^{10-j}$$

Generalizing to arbitrary n and k,

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

Does this generalize for us?

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