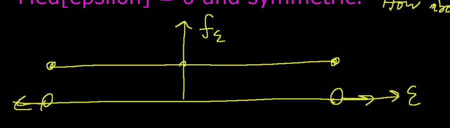
1774. First "law of errors". Imagine you're trying to measure something, a const quantity v, but your measurements have random error, epsilon, so your measurement M is a rv looking like: M = v + epsilon. So what is a good model for the error (epsilon)? It makes sense for E[epsilon] = 0. Med[epsilon] = 0 and symmetric. How about... $U(-L, L)$



It also makes sense for larger errors (in magnitude) to be less probable than smaller errors. $\Rightarrow \forall \varepsilon > 0 \quad f'(\varepsilon) < 0$

$$\forall \varepsilon > 0 \quad \overset{\text{If...}}{f''(\varepsilon) = f'(\varepsilon)} \underset{\text{solve}}{\Rightarrow} f(\varepsilon) = c e^{-d\varepsilon} \Rightarrow \text{Laplace}(0, 1)$$

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

$$Y \sim f_Y(y) = ? \quad \text{Inverse function} \quad \lambda y = x^{\frac{1}{k}} \Rightarrow x = \lambda^k y^k = (\lambda y)^k = g^{-1}(y)$$

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

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

$$= k\lambda^k y^{k-1} e^{-(\lambda y)^k} \mathbb{1}_{y \geq 0} = k\lambda(\lambda y)^{k-1} e^{-(\lambda y)^k} \mathbb{1}_{y \geq 0} = \text{Weibull}(k, \lambda) \quad 1951$$

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

k is really cool... this is the main property: e.g. $y = 3$, $c = 14$

$$k = 1 \quad P(Y \geq y + c \mid Y \geq c) = P(Y \geq y). \qquad P(Y \geq 17 \mid Y \geq 14) = P(Y \geq 3)$$

$\rightarrow$ Memorylessness

$$k > 1 \quad P(Y \geq y + c \mid Y \geq c) < P(Y \geq y)$$ $\Rightarrow$ survival less likely as time goes on

$$k < 1 \quad P(Y \geq y + c \mid Y \geq c) > P(Y \geq y)$$ $\Rightarrow$ survival more likely as time goes on

you will prove these facts on the HW

Order Statistics (p160 in the textbook).

Let $X_1, X_2, \ldots, X_n$ be a collection of continuous rv's and let

$X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be their "order statistics" defined as:

$X_{(1)} = \min\{X_1, \ldots X_n\}$ minimum

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

$\vdots$

$X_{(k)} = k^{th}$ largest $\{X_1, \ldots, X_n\}$

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

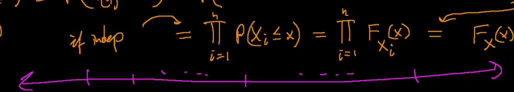
$x_1 = 9, x_2 = 2, x_3 = 12, x_4 = 7$

$x_{(1)} = 2, x_{(2)} = 7, x_{(3)} = 9, x_{(4)} = 12$

$\Rightarrow r = 12 - 2 = 10.$

We want to find the CDF and PDF of the order statistics. We'll start by looking at the CDF/PDF of the maximum.

$$F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = P(X_1 \leq x \;\&\; X_2 \leq x \;\&\; \ldots \;\&\; X_n \leq x) \overset{\text{if indep}}{=} \prod_{i=1}^n P(X_i \leq x) = \prod_{i=1}^n F_{X_i}(x) \overset{\text{if iid}}{=} F_X(x)^n$$



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

Let's now find the CDF/PDF of the minimum.

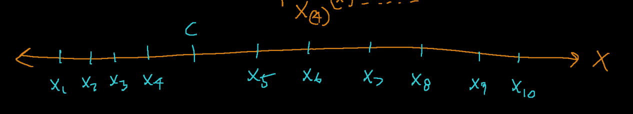
$$F_{X_{(1)}}(x) = P(X_{(1)} \leq x) = 1 - P(X_{(1)} > x) = 1 - P(X_1 > x \;\&\; X_2 > x \;\&\; \ldots \;\&\; X_n > x)$$

$$\overset{\text{if indep}}{=} 1 - P(X_1 > x) P(X_2 > x) \cdots P(X_n > x) = 1 - \prod_{i=1}^n \left(1 - F_{X_i}(x)\right)$$

$$\overset{\text{if iid}}{=} 1 - (1 - F_X(x))^n$$

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

Let's now find the CDF/PDF for the kth order statistic, $X_{(k)}$.
Let's let n = 10, k = 4. $F_{X_{(4)}}(x) = \ldots$



$$P(X_1 \leq c \;\&\; \ldots \;\&\; X_4 \leq c \;\&\; X_5 > c \;\&\; \ldots \;\&\; X_{10} > c)$$

$$\overset{\text{if indep}}{=} \prod_{i=1}^4 P(X_i \leq c) \prod_{i=5}^{10} P(X_i > c) = \prod_{i=1}^4 F_{X_i}(c) \prod_{i=5}^{10} (1 - F_{X_i}(c)) \overset{\text{if iid}}{=} F_X(c)^4 (1 - F_X(c))^6$$

$$P(\text{any } 4\ X_i\text{'s} \leq x \;\&\; \text{the other } 6\ x_i\text{'s} > x)$$

$$= \sum_{\substack{\text{over all subsets} \\ S \text{ size } 4,\ S^c \\ \text{size } 6}} P(X_{S_1} \leq x, \ldots, X_{S_4} \leq x, X_{S^c_1} > x, \ldots, X_{S^c_6} > x)$$

$$\overset{\text{if indep}}{=} \sum_{\text{same}} \prod_{i=1}^4 F_{X_{S_i}}(x) \prod_{i=1}^6 F_{X_{S^c_i}}(x) \overset{\text{if iid}}{=} \sum_{\text{same}} F_X(x)^4 (1 - F_X(x))^6$$

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

$$F_{X_{(4)}}(x) = P(X_{(4)} \leq x) = P(4\ X_i\text{'s} \leq x, 6\ X_i\text{'s} > x) + P(5\ X_i\text{'s} \leq x, 5\ X_i\text{'s} > x) + \ldots + P(10\ X_i\text{'s} \leq x, 0\ X_i\text{'s} > x)$$

$$\cancel{+ P(\underbrace{3\ X_i\text{'s} \leq x, 7\ X_i\text{'s} > x}_{X_{(4)} > x})}$$

$$\overset{\text{if iid}}{=} \sum_{j=4}^{10} \binom{10}{j} F_X(x)^j (1 - F_X(x))^{10-j}$$

general case: k, n

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

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

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

Binomial Thm.: $(a+b)^n = \sum_{j=0}^n \binom{n}{j} a^j b^{n-j}$

$$= \left(\sum_{j=0}^n \binom{n}{j} \underbrace{F_X(x)}_{a}{}^j \underbrace{(1 - F_X(x))}_{b}{}^{n-j}\right) - \binom{n}{0} F_X(x)^0 (1 - F_X(x))^{n-0}$$

$$= (F_X(x) + 1 - F_X(x))^n - (1 - F_X(x))^n = 1 - (1 - F_X(x))^n$$

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

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

$$\frac{d}{dx}[uv] = uv' + u'v \qquad u' = j f(x) F(x)^{j-1} \qquad v' = -(n-j) f(x) (1 - F(x))^{n-j-1}$$