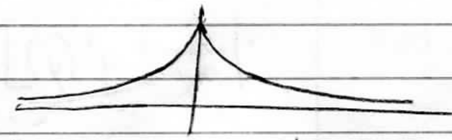
$D \sim \text{Laplace}(0,1) = \frac{1}{2}e^{-|d|}$



Laplace discoved this model in 1774 calling it the first "law of error"

Imagine you are measuring a quantity but your measurement device has some random additive error, ε.
So your measurement is $M = v + \varepsilon$.
Thus M is random
actual value

Certain properties of $\varepsilon \sim f(\varepsilon)$:
pdf

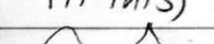
★ $E[\varepsilon] = 0 \Rightarrow E[M] = v$
(this means that M is an "unbiased estimator")

★ $Med[E] = 0 \Rightarrow$ 50% of the time you overestimate
50% of the time you underestimate

★ $f(\varepsilon) = f(-\varepsilon)$ symmetric

★ $f'(\varepsilon) < 0$ if $\varepsilon > 0$ and $f'(\varepsilon) > 0$ if $\varepsilon < 0$
(the further away from 0, the smaller the probabilities)

this is general concept of an error distribution (many different distributions fit this)



Consider $f''(\varepsilon) = f'(\varepsilon) \Rightarrow f(\varepsilon) = ce^{-b\varepsilon} = \text{Laplace}(0,1)$

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

$\Rightarrow \lambda Y = X^{\frac{1}{k}}$

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

$Y = X^k$
$Y = X^{\alpha}$

$X \sim Exp(1) = e^{-x}\,\mathbb{1}_{x \geq 0}$

Find pdf of
$Y = \frac{1}{\lambda} X^{\frac{1}{k}}$
$g(X) = \frac{1}{\lambda} X^{\frac{1}{k}}$

got $g^{-1}(Y) = \lambda^k Y^k$

Weibull is common survival model

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

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

$y^k \geq 0$
$y \geq 0$

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

$= Weibull(k, \lambda)$

Very functional waiting time/survival model

Weibull is a generalization of the Exponential
$Weibull(k=1, \lambda) = \lambda e^{-\lambda y} = Exp(\lambda)$

CDF
$F(y) = \int_0^y (k\lambda)(\lambda t)^{k-1} e^{-(\lambda t)^k}\, dt =$

let $u = (\lambda t)^k$
$\frac{du}{dt} = \lambda^k k t^{k-1}$
$dt = \frac{1}{\lambda^k k t^{k-1}}\, du$

$t = y \rightarrow u = (\lambda y)^k$
$t = 0 \rightarrow y = 0$

$\bar{F}(y)$ survival function since represents $P(Y > y)$ (last longer than y)

$= \int_0^{(\lambda y)^k} k \cancel{\lambda^k t^{k-1}} e^{-u} \frac{1}{\cancel{\lambda^k k t^{k-1}}}\, du = \int_0^{(\lambda y)^k} e^{-u}\, du$

$= \left[-e^{-u}\right]_0^{(\lambda k)^y} = -\left(e^{-(\lambda k)^y} - e^0\right) =$

$\bar{F}(y) = 1 - F(y)$
$\bar{F}(y) = e^{-(\lambda y)^k}$

$F(y) = 1 - e^{-(\lambda y)^k}$ ← CDF for Weibull survival function $\overline{F}(y) = 1 - F(y) = e^{-(\lambda y)^K}$

Is the Weibull memoryless?

memoryless means tha prob. that it takes c time units more is the same as prob. that takes c years

does not matter how much time passed already is "Memoryless"

$$P(Y \geq y + c \mid Y \geq c) = \frac{P(Y \geq y+c,\ Y \geq c)}{P(Y \geq c)}$$

$$= \frac{P(Y \geq y+c)}{P(Y \geq c)}$$

$$= \frac{\overline{F}(y+c)}{\overline{F}(c)} = \frac{e^{-(\lambda(y+c))^K}}{e^{-(\lambda c)^K}}$$

by weibull

$$= e^{-\lambda^K (y+c)^K} e^{\lambda^K c^K}$$

$$= e^{\lambda^K (c^K - (y+c)^K)}$$

If $K = 1$ (this covers Exponential)

$$P(Y \geq y + c \mid Y \geq c) = e^{\lambda(c - (y+c))}$$

$$= e^{-\lambda y} =$$

$$= P(Y \geq y)$$

Weibull is memoryless only for $K=1$ (Expoental)

this is Memoryless (meaning $P(Y \geq y+c \mid Y \geq c) = P(Y \geq y)$)

but if $K > 1$:

$$P(Y \geq y + c \mid Y \geq c) < P(Y \geq y)$$

this inequality gets more severe as we increase K

human lifespan

Ex: $P(Y \geq 100 \mid Y \geq 99) < P(Y \geq 1)$

prob. of lasting an additional year < prob. lasts a year inequality more severe as decrease K

If $K<1 \Rightarrow P(Y \geq y+c \mid Y \geq c) > P(Y \geq y)$

infant mortality, "burn-in" period for machines or electronics

$P(Y \geq 15 \text{ years an } 1 \text{ day} \mid Y \geq 15 \text{ years}) \leq P(Y \geq 1 \text{ day})$

prob. of lasting an additional day is less than prob. of lasting a day

Verify the inequality if $K=2$

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

$$\lambda^2(c^2-(y-c)^2) < -\lambda^2 y^2$$

$$c^2+y^2 < (y+c)^2$$

$$c^2+y^2 < y^2+2yc+c^2$$

$$0 < 2yc$$

since $yc > 0$ ✓

Verify the inequality if $K=\frac{1}{2}$

$$c^{\frac{1}{2}} + y^{\frac{1}{2}} > (y+c)^{\frac{1}{2}}$$

$$\left[c^{\frac{1}{2}} + y^{\frac{1}{2}}\right]^2 > \left[(y+c)^{\frac{1}{2}}\right]^2$$ square both sides

$$c + 2c^{\frac{1}{2}}y^{\frac{1}{2}} + y > y + c$$

$$c + 2c^{\frac{1}{2}}y^{\frac{1}{2}} > c$$

$$2c^{\frac{1}{2}}y^{\frac{1}{2}} > 0$$ ✓

because $c>0,\ y>0$

so $c^{\frac{1}{2}}>0,\ y^{\frac{1}{2}}>0$

so $c^{\frac{1}{2}}y^{\frac{1}{2}}>0$

so $2c^{\frac{1}{2}}y^{\frac{1}{2}}>0$

note that left 5th one out

Consider $P($4 of the 10 r.v.'s $\le x$
and the remaining 6 of the 10 r.v.'s $> x)$

actually $\sum_{\{S_1, S_2, S_3, S_4\} \subset \{1, 2, \dots 10\}}$

$$= \sum P(X_{S_1} \le x, \dots, X_{S_4} \le x, X_{S_5} > x, \dots X_{S_{10}} \geq x)$$

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

since are i.i.d.,

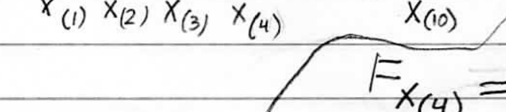
$$= \sum (F(x))^4 (1 - F(x))^6$$

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

$$F_{X_{(4)}} = P(X_{(4)} \le x) \underset{\text{if i.i.d.}}{=} \binom{10}{4} F(x)^4 (1 - F(x))^6 + \binom{10}{5} (F(x))^5 (1 - F(x))^5 + \binom{10}{6} (F(x))^6 (1 - F(x))^4 + \dots + \binom{10}{10} (F(x))^{10} (1 - F(x))^0$$

Ex: if 4th largest is ≤ 3.7

3.7 could be here



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

verify for $F_{X_{(n)}}, F_{X_{(1)}}$

here, had $k = 4$, $n = 10$

In general,

For arbitrary n, k, and r.v.'s $X_1, X_2, \dots, X_{10}$ which are i.i.d.

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

use binomial CDF? $p = F(x)$

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

CDF of max

$$= \binom{n}{n} (F(x))^n (F(x))^{n-n} \leftarrow (F(x))^0 = 1$$

$$= 1 \; (F(x))^n \; 1$$

$$F_{X_{(n)}}(x) = (F(x))^n$$

using

$$\sum_{K=0}^{n} K(i) = K(0) + \sum_{K=1}^{n} K(i)$$

or $\sum_{K=1}^{n} K(i) = \sum_{k=0}^{n} K(i) - K(0)$

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

CDF of min

$$= \left(\sum_{j=0}^{n} \binom{n}{j} (F(x))^j (1-F(x))^{n-j} \right) - \binom{n}{0} (F(x))^0 (1-F(x))^n$$

this is a sum of all the binomial probobabilites so must be 1

$$\sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} = 1$$

$$= 1 - 1 \cdot 1 \cdot (1-F(x))^n$$

$$F_{X_{(1)}} = 1 - (1-F(x))^n$$

for this summation, notice

$$(F(x) + (1-F(x)))^n$$

$$= (1)^n$$

$$= 1$$

so $F_{X_{(1)}} = 1 - (1-F(x))^n$ ← CDF of min

In general,

For arbitrary n, K and i.i.d r.v.'s $X_1, X_2, \ldots, X_n$

CDF of K th largest one

CDF

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

PDF

$$f_{X_{(K)}}(x) = \frac{d}{dx}\left[F_{X_{(K)}}(x)\right] = \frac{d}{dx}\left[\text{all of that stuff}\right] =$$

done in class next time

Order Statistics (p.160-161)

Consider continuous r.v.'s $X_1, X_2, \ldots, X_n$

Define $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ as the "order statistics" as follows:

$X_{(1)} := \min\{X_1, \ldots, X_n\}$ $\quad X_{(n)} := \max\{X_1, \ldots, X_n\}$

$X_{(k)}$ is kth one if we put in order smallest to largest

$X_{(k)}$ is the kth largest of $\{X_1, \ldots, X_n\}$ called "k^{th} order statistics"

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

Example

For $n=4$ realizations

$x_1 = 9,\ x_2 = 2,\ x_3 = 7,\ x_4 = 14$

$x_{(1)} = 2,\ x_{(2)} = 7,\ x_{(3)} = 9,\ x_{(4)} = 14$

$r = 19 - 2 = 12$

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

$$
\begin{aligned}
F_{X_{(n)}}(x) &:= P(X_{(n)} \le x) \\
&= P(X_1 \le x, X_2 \le x, \ldots, X_n \le x) \\
&\text{if } X_1, X_2, \ldots, X_n \text{ are independent} \\
&= P(X_1 \le x) P(X_2 \le x) \cdots P(X_n \le x) \\
&\text{if } X_1, X_2, \ldots X_n \text{ are i.i.d.} \\
&= \underbrace{P(X \le x) P(X \le x) \cdots P(X \le x)}_{n \text{ of these}} \\
&= [P(X \le x)]^n
\end{aligned}
$$

$$F_{X_{(n)}}(x) = [F(x)]^n$$

${}^{n}F'(x)$

CDF $F_{X_{(n)}}(x) = [F(x)]^n$

to get PDF do derivative

since X's are i.i.d. $\boxed{F'(x) = f(x)}$

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

---

Now, look at the minimum $X_{(1)}$

CDF

$$F_{X_{(1)}}(x) = P(X_{(1)} \le x) = 1 - P(X_{(1)} > x)$$

$$= 1 - P(X_1 > x, X_2 > x, \dots, X_n > x)$$

if are independent

$$= 1 - P(X_1 > x)P(X_2 > x) \cdots P(X_n > x)$$

if are i.i.d.

$$= 1 - \underbrace{P(X > x)P(X > x) \cdots P(X > x)}_{n \text{ of these}}$$

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

$$F_{X_{(1)}}(x) = 1 - [F(X)]^n$$

find PDF by taking the derivative

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

$$= n f(x) (1 - F(x))^{n-1}$$

Let's find CDF of $X_{(K)}$, the k^{th} ordered statistic

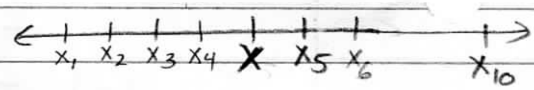
Consider $n = 10$, $k = 4$

$$F_{X_{(4)}}(x) = P(X_{(4)} \le x) \quad \text{Assume } X\text{'s are i.i.d.}$$

$$= P(X_1 \le x, X_2 \le x, X_3 \le x, X_4 \le x, X_5 > x, X_6 > x, \dots, X_{10} > x)$$

$$= \prod_{i=1}^{5} F_x(x) \prod_{i=6}^{10} (1 - F_x(x))$$

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



this works if all of them are in order