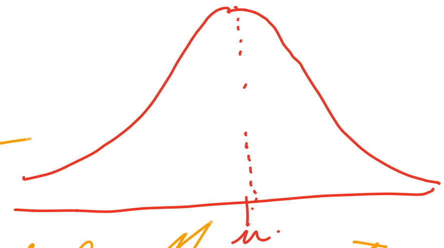
Lecture 23



Last time: If X is exponential with parameter λ, and $a>0$, then the variable

$$Y = a e^{X}$$

is Pareto with parameters a, λ. If Y is such, then

$$F_Y(y) = 1 - a^{\lambda} y^{-\lambda}, \quad y \geq a$$

and

$$f_Y(y) = \lambda a^{\lambda} y^{-(\lambda+1)}, \quad y \geq a.$$

Lets compute $E[Y]$:

$$E[Y] = \int_a^{\infty} \lambda a^{\lambda} y \, y^{-\lambda-1} \, dy.$$

$$= \lambda a^{\lambda} \int_a^{\infty} y^{-\lambda} \, dy = \lambda a^{\lambda} \left[\frac{y^{1-\lambda}}{1-\lambda} \right]_a^{\infty}$$

$$= \lambda a^{\lambda} \left(\frac{a^{1-\lambda}}{\lambda - 1} + \lim_{t \to \infty} \frac{t^{1-\lambda}}{1-\lambda} \right.$$

Now,

$$\lim_{t \to \infty} \frac{t^{1-\lambda}}{1-\lambda} = \begin{cases} \infty & \text{if } \lambda \leq 1 \\ 0 & \text{if } \lambda > 1 \end{cases}$$

Therefore $$E[Y] = \begin{cases} \infty & \text{if } \lambda \le 1 \\ \dfrac{\lambda a}{\lambda - 1} & \text{if } \lambda > 1. \end{cases}$$

By the same type of integral,

$$E[Y^2] = \int_a^\infty \lambda a^\lambda y^2 y^{-\lambda-1} dy$$
$$= \begin{cases} \infty & \text{if } \lambda \le 2 \\ \dfrac{\lambda a^2}{\lambda-2} & \text{if } \lambda > 2. \end{cases}$$

So $$\mathrm{Var}(Y) = \frac{\lambda a^2}{\lambda-2} - \frac{\lambda^2 a^2}{(\lambda-1)^2} = \frac{\lambda a^2}{(\lambda-2)(\lambda-1)^2}.$$

Note that we can also compute the moments as

$$E[Y^n] = E[a^n e^{nX}] = a^n E[e^{nX}] = a^n \int_0^\infty e^{nx} \lambda e^{-\lambda x} dx$$
$$= \lambda a^n \int_0^\infty e^{(n-\lambda)x} dx = \lambda a^n \left(\frac{e^{(n-\lambda)x}}{n-\lambda} \Big|_0^\infty \right)$$
$$= \begin{cases} \infty & \text{if } \lambda \le n \\ \dfrac{\lambda a^n}{\lambda - n} & \text{if } \lambda > n. \end{cases}$$

The pareto distribution can be used to model:

- income/wealth of members of a population
- file size of internet traffic.
- time to complete a job assigned to a super

computer.

- the size of a meteorite.
- yearly maximum 1 day rainfalls in a given region.

The distribution of a function of a RV.

- Given a RV X, often want to determine the distribution of $Y = g(X)$ for some function g.
- to do this, we want to be able to express the event $g(X) \le y$ in terms of X being in some set.

Ex: Let X be a CRV with density f_X. If $Y = X^2$, then, for $y \geq 0$:

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(X^2 \le y) \\ &= P(-\sqrt{y} \le X \le \sqrt{y}) \leftarrow \text{this is what we want} \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

Differentiating gives:

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right].$$

Ex: Suppose X is a CRV with density f_X.
Let $Y = |X|$. Find f_Y:

$$F_Y(y) = P(Y \le y) = P(|X| \le y)$$
$$= P(-y \le X \le y)$$
$$= F_X(y) - F_X(-y).$$

So $$f_Y(y) = f_X(y) + f_X(-y), \quad y \geq 0.$$

We have a general theorem:

Theorem: Let X be a CRV with pdf f_X.
Suppose that $g(x)$ is strictly monotonic and differentiable. Then $Y = g(X)$ has pdf

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| & \text{if } y \in \text{range}(g) \\ 0 & \text{if } y \notin \text{range}(g). \end{cases}$$

Proof: We prove the theorem in the case where $g(x)$ is increasing. Note that strictly monotonic functions are 1-1, so g^{-1} is well-defined.

If $Y = g(X)$, then

$$F_Y(y) = P(g(X) \le y)$$
$$= P(X \le g^{-1}(y)) \quad \text{(for all } y \in \text{range}(X)\text{)}$$
$$= F_X(g^{-1}(y)).$$

The chain rule gives:

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \underbrace{\frac{d}{dy}\left(g^{-1}(y)\right)}_{\ge 0}$$

≥ 0, since $g^{-1}(y)$ is nondecreasing

Suppose $y \in \text{Range}(g)$. Then either

$y < g(x)$ $\forall x \in \text{dom}(g)$ in which case $F_Y(y) = 0$.

$y > g(x)$ for every $x \in \text{dom}(g)$ in which case $F_Y(y) = 1$.

either way, $f_Y(y) = 0$, as required.

Ex: The lognormal distribution.

Let $X = N(\mu, \sigma^2)$, the normal distribution with mean μ and variance σ^2.

Then the CRV $Y = e^X$ is called lognormal.

The lognormal distribution is used in securities pricing.

Recall that $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$.

So if $y = e^X = g(X)$, then

$$g^{-1}(y) = \log(y).$$

By the theorem,

$$f_Y(y) = \frac{e^{-(\log(y)-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \cdot \frac{d}{dy}\left(\log(y)\right)$$

$$= \frac{e^{-(\log(y)-\mu)^2/2\sigma^2}}{\sigma y\sqrt{2\pi}}.$$