# Scale Distributions [cont'd].

M339J: February 12th, 2020.

**Example.** Start w/

$$X \sim \text{Normal}(\text{mean} = \mu, \text{var} = \sigma^2)$$

Set

$$Y = e^X$$

We say that $Y$ has the

**Lognormal distribution.**

We choose $\mu$ and $\sigma$ as the parameters of the lognormal dist'n as well.

Q: What is the c.d.f. of $Y$?

The support of $Y$ is the positive half-line.

For $y > 0$:

$$F_Y(y) = \mathbb{P}[Y \leq y] = \mathbb{P}[e^X \leq y]$$

ln is strictly increasing

$$= \mathbb{P}[X \leq \ln(y)]$$

$$= \mathbb{P}\Big[\underbrace{\frac{X-\mu}{\sigma}}_{\sim N(0,1) \sim Z} \leq \frac{\ln(y)-\mu}{\sigma}\Big]$$

$$= \mathbb{P}\Big[Z \leq \frac{\ln(y)-\mu}{\sigma}\Big]$$

$$= \Phi\Big(\frac{\ln(y)-\mu}{\sigma}\Big) \qquad \square$$

Q: Is the lognormal a scale dist'n?

Consider a constant $c > 0$

Set $\tilde{Y} = c \cdot Y$

Look @ $\tilde{Y} = c \cdot Y = c \cdot e^X = e^{\overbrace{\ln(c) + X}^{\tilde{X}}}$

**Note:** $\tilde{X} \sim \text{Normal}(\text{mean} = \mu + \ln(c), \text{var} = \sigma^2)$

$\Rightarrow \tilde{Y}$ is also lognormal

The lognormal is a scale dist'n, but it does not have a scale parameter in this parametrization.

## k-point mixture

A random variable $Y$ is a k-point mixture of the random variables $X_1, X_2, \ldots, X_k$ if its c.d.f. is given by

$$F_Y(y) = a_1 \cdot F_{X_1}(y) + a_2 \cdot F_{X_2}(y) + \cdots + a_k \cdot F_{X_k}(y)$$

where $a_j$, $j = 1..k$, are positive constants such that $a_1 + a_2 + \cdots + a_k = 1$

**168.** For an insurance:

(i) Losses can be 100, 200 or 300 with respective probabilities 0.2, 0.2, and 0.6.

(ii) The insurance has an ordinary deductible of 150 per loss.

(iii) $Y^P$ is the claim payment per payment random variable.

Calculate $\mathrm{Var}(Y^P)$.

(A) 1500

(B) 1875

(C) 2250

(D) 2625

(E) 3000

***169.** The distribution of a loss, *X*, is a two-point mixture:

$a_1 = 0.8$

(i) With probability 0.8, *X* has a two-parameter Pareto distribution with $\alpha = 2$ and $\theta = 100$.

$a_2 = 0.2$

(ii) With probability 0.2, *X* has a two-parameter Pareto distribution with with $\alpha = 4$ and $\theta = 3000$.

Calculate $\Pr(X \leq 200)$.

By def'n $\mathbb{P}[X \leq 200] = F_X(200)$

(A) 0.76

(B) 0.79

$$X \sim \begin{cases} X_1 \sim \text{Pareto}(\alpha_1 = 2, \theta_1 = 100) & \text{w/ prob. } a_1 = 0.8 \\ X_2 \sim \text{Pareto}(\alpha_2 = 4, \theta_2 = 3000) & \text{w/ prob. } a_2 = 0.2 \end{cases}$$

(C) 0.82

(D) 0.85

(E) 0.88

4.

Since X is a 2-point mixture, its cdf is

$$F_X(x) = a_1 \cdot F_{X_1}(x) + a_2 \cdot F_{X_2}(x)$$
$$= 0.8\left[1-\left(\frac{100}{x+100}\right)^2\right] + 0.2\left[1-\left(\frac{3000}{x+3000}\right)^4\right]$$
$$= 1 - 0.8\left(\frac{100}{x+100}\right)^2 - 0.2\left(\frac{3000}{x+3000}\right)^4$$

Answer :

$$F_X(200) = 1 - 0.8\left(\frac{100}{300}\right)^2 - 0.2\left(\frac{3000}{3200}\right)^4$$
$$= 1 - \frac{4}{5}\cdot\frac{1}{9} - \frac{1}{5}\cdot\left(\frac{15}{16}\right)^4 = 0.7566$$

$\Rightarrow$ (A)

**287.** For an aggregate loss distribution $S$:

(i) The number of claims has a negative binomial distribution with $r = 16$ and $\beta = 6$.

(ii) The claim amounts are uniformly distributed on the interval (0, 8).

(iii) The number of claims and claim amounts are mutually independent.

Using the normal approximation for aggregate losses, calculate the premium such that the probability that aggregate losses will exceed the premium is 5%.

(A) 500

(B) 520

(C) 540

(D) 560

(E) 580

* **288.** The random variable $N$ has a mixed distribution: probab. of success # of trials

$a_1$

(i) With probability $p$, $N$ has a binomial distribution with $q = 0.5$ and $m = 2$.

$a_2$

(ii) With probability $1-p$, $N$ has a binomial distribution with $q = 0.5$ and $m = 4$.

Which of the following is a correct expression for $\Pr(N = 2)$?

(A) $0.125p^2$

(B) $0.375 + 0.125p$

(C) $0.375 + 0.125p^2$

(D) $0.375 - 0.125p^2$

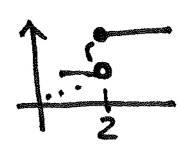
(E) $0.375 - 0.125p$

$$N \sim \begin{cases} N_1 \sim \text{bin}(q_1 = \tfrac{1}{2}, m_1 = 2) & \text{w/ probab. } p \\ N_2 \sim \text{bin}(q_2 = \tfrac{1}{2}, m_2 = 4) & \text{w/ probab. } 1-p \end{cases}$$

We seek

$$p_N(2) = \mathbb{P}[N=2] = F_N(2) - F_N(2-)$$

$$= F_N(2) - F_N(1)$$




By $N$ being a two-point mixture

$$p_N(2) = p \cdot F_{N_1}(2) + (1-p) \cdot F_{N_2}(2)$$

$$- \left(p \cdot F_{N_1}(1) + (1-p) \cdot F_{N_2}(1)\right)$$

$$\Rightarrow p_N(2) = p\,\underbrace{\left(F_{N_1}(2) - F_{N_1}(1)\right)}_{= p_{N_1}(2)} + (1-p)\underbrace{\left(F_{N_2}(2) - F_{N_2}(1)\right)}_{= p_{N_2}(2)}$$

$$\Rightarrow p_N(2) = p \cdot p_{N_1}(2) + (1-p) \cdot p_{N_2}(2)$$

<u>In general</u>:

Say $N_1, N_2, \ldots, N_k$ are $\mathbb{N}_0$-valued random variables. Let $p_{N_i}(\cdot)$ denote the p.m.f.s of $N_i$, $i = 1..k$.

Define the k-pt mixture $N$ as

$$N \sim \begin{cases} N_1 & \text{w/ prob. } a_1 \\ N_2 & \text{w/ prob. } a_2 \\ \vdots & \\ N_k & \text{w/ prob. } a_k \end{cases} \qquad \sum a_i = 1, \quad a_i > 0$$

<u>Then</u>, for any $j \in \mathbb{N}_0$, we have that

$$p_N(j) = a_1 \cdot p_{N_1}(j) + a_2 \cdot p_{N_2}(j) + \cdots + a_k \cdot p_{N_k}(j)$$

Returning to our problem, we get

$$\mathbb{P}[N=2] = p \cdot \left(\tfrac{1}{2}\right)^2 + (1-p)\binom{4}{2}\left(\tfrac{1}{2}\right)^4$$

$$= 0.25p + (1-p) \cdot \frac{4 \cdot 3}{2} \cdot \frac{1}{16}$$

$$= 0.25p + (1-p)(0.375)$$

$$= 0.375 - 0.125p \Rightarrow (E) \blacksquare$$