Creating New Distributions

Section 5.2

Stat 346 - Short-term Actuarial Math

Some methods to generate new distributions

- There are many methods to generate new distributions; some of these methods allow us to give in-depth interpretation to the distributions.
- Among the methods used can be sub-divided into:
  1. Addition of several random variables
     - For example, sums of (independent) Exponentials give a Gamma. This method will not be further explored.
  2. Transformation of random variables
     - scalar multiplication
     - power operations
     - exponentiation (or logarithmic transformation)
  3. Mixing of distributions
     - frailty models
  4. Spliced distributions
- Section 5.2 of Klugman, et al.

The general theory of transformation

Suppose we are interested in deriving the distribution of $Y = g(X)$, where X has a known distribution function. Assume that the function g is a one-to-one transformation (i.e. invertible).

- It can be shown that the distribution function of Y can be expressed as

  $$F_Y(y) = F_X(g^{-1}(y)),$$

  in the case of increasing transformation. If decreasing, we have

  $$F_Y(y) = 1 - F_X(g^{-1}(y)).$$

- Its density can be explicitly written as

  $$f_Y(y) = f_X(g^{-1}(y)) \times \left| \frac{dg^{-1}(y)}{dy} \right|.$$

Scalar transformations

- In the case where $Y = aX$ for some $a > 0$, then this is called a scalar transformation and its density function can be expressed as

$$f_Y(y) = \frac{1}{a} f_X(y/a).$$

- Insurance interpretation: if X denotes claims, then scalar transformation can be interpreted as applying inflation factor across all levels of claims.
- A family of distributions that is closed under scalar multiplication (i.e. after scalar transformation, the new random variable remains in the same family) is called a scale family of distributions.
- Some scale families are:
  - Normal
  - Exponential (Example 5.1)
  - Pareto

Power transformations

A power transformation involves raising the random variable by a power such as

$$Y = X^{1/\tau} \text{ or } Y = X^{-1/\tau},$$

where $\tau > 0$.

In the first case, we have a transformed X distribution; the other case, we have an inverse transformed X distribution.

In the special case where $Y = X^{-1}$, we have an inverse X distribution.

Distribution and density functions of power transformations

It is easy to show the following results:

- In the transformed case where $Y = X^{1/\tau}$, we have

$$F_Y(y) = F_X(y^\tau) \text{ and } f_Y(y) = \tau y^{\tau-1} f_X(y^\tau).$$

- In the inverse transformed case where $Y = X^{-1/\tau}$, we have

$$F_Y(y) = 1 - F_X(y^{-\tau}) \text{ and } f_Y(y) = \tau y^{-\tau-1} f_X(y^{-\tau}).$$

- In the inverse case where $Y = X^{-1}$, we have

$$F_Y(y) = 1 - F_X(y^{-1}) \text{ and } f_Y(y) = \frac{1}{y^2} f_X(1/y).$$

Example 5.2

Let X be exponentially distributed with mean parameter 1.

Derive the cumulative distribution and density functions of the transformed, inverse transformed and inverse random variables.

Note that we derive:

- Inverse Exponential distribution: for the case of the inverse distribution after a scale transformation
- Weibull distribution: for the case of the transformed distribution after a scale transformation
- Inverse Weibull distribution: for the case of the inverse transformed distribution after a scale transformation

Illustrative example 1

Suppose X is a random variable with cumulative distribution function

$$F_X(x) = 1 - (1 + x^c)^{-\gamma}, \quad x > 0, \quad c > 0, \quad \gamma > 0.$$

Derive the probability density function of the inverse of X, i.e. $Y = 1/X$.

Illustrative example 2

Suppose X is an exponential random variable with mean parameter equal to 1.

Derive the distribution of $Y = \alpha X^{1/\beta}$, for $\alpha > 0, \beta > 0$. Specify its density function. The distribution of Y is called a Weibull and we can write $Y \sim$ Weibull(α, β), where α is obviously a scale parameter.

Exponentiation

Suppose $Y = \exp(X)$. For $y > 0$, its distribution function can be expressed as

$$F_Y(y) = F_X(\log y)$$

and its corresponding density as

$$f_Y(y) = \frac{1}{y} f_X(\log y)$$

Derive the distribution/density functions corresponding to the exponential transformation $Y = \exp(X)$ when:

- $X \sim \mathsf{N}(\mu, \sigma^2)$
- $X \sim \mathsf{Exp}(1)$

Mixtures of distributions

A random variable X is said to be a mixture of distributions if its distribution function has one of the following forms:

1. Discrete mixture: $F_X(x) = \sum_i a_i F_{X_i}(x)$, for some sequence of random variables $X_1, X_2, \ldots$ and some sequence of positive numbers $a_1, a_2, \ldots$ satisfying $\sum_i a_i = 1$.
2. Continuous mixture: $F_X(x) = \int_{\infty}^{-\infty} F_{X|\Lambda=\lambda}(x) f_\Lambda(\lambda) d\lambda$, for some random variable Λ satisfying $\int_{\infty}^{-\infty} f_\Lambda(\lambda) d\lambda = 1$.

In terms of actuarial/insurance applications:

- Discrete mixtures arise in situations where the risk class of a policyholder is uncertain, and the number of possible risk classes is discrete.
- Continuous mixtures arise when a risk parameter from the loss distribution is uncertain and the uncertain parameter is continuous.

Illustrative example 1

An insurer has two groups of policyholders: the good and the bad risks. The insurer has a portfolio where 75% are considered good risks.

The claim distributions for both groups of risks are Exponential. The average claim amount of a good risk policyholder is \$100 while for bad risk, it is twice that.

A new customer whose risk class is not known with certainty, has just recently purchased a policy from the insurer.

Calculate the probability that this new customer will claim an amount exceeding \$150.

Illustrative example 2

Consider a claims random variable X that, given a risk classification (random) parameter Λ, can be modeled as an Exponential random variable with

$$P(X \leq x|\Lambda = \lambda) = 1 - e^{-\lambda x}, \quad \text{for } x > 0.$$

Assume that Λ has a Gamma$(\alpha, 1/\theta)$ distribution.

Show that the unconditional distribution of X is a Pareto.

Deriving the mean and variance of mixtures

Suppose that X is a mixture with mixing variable Λ. Then the unconditional mean and variance can be determined using the following formulas:

- Law of iterated expectations

$$\mathsf{E}(X) = \mathsf{E}_\Lambda[\mathsf{E}(X|\Lambda)],$$

  or in general, we have

$$\mathsf{E}(X^k) = \mathsf{E}_\Lambda[\mathsf{E}(X^k|\Lambda)].$$

- Conditional variance formula

$$\mathsf{Var}(X) = \mathsf{E}_\Lambda[\mathsf{Var}(X|\Lambda)] + \mathsf{Var}_\Lambda[\mathsf{E}(X|\Lambda)].$$

Illustration

The policyholders of an insurance company fall into one of two classes. The claims distributions for each class are given in the following table:

Class 1		Class 2	
claim size	probability	claim size	probability
1,000	0.20	1,000	0.70
5,000	0.50	5,000	0.20
10,000	0.30	10,000	0.10

There are 30% of policyholders in class 1, while the remaining policyholders are in class 2. Denote by L_1 the claim incurred by a randomly selected policyholder from class 1, while L_2 from class 2.

Let L denote the claim incurred by a randomly selected policyholder whose risk class is unknown.

Calculate the mean and variance of L.