Shyamal Kumar

Mixed Type Distributions

Shyamal Kumar

1 Preface

It is common, for many good reasons, that elementary probability courses do not cover distribution of the mixed type - by mixed type we mean distributions with a discrete as well as a continuous component. This implies that many actuarial science students embark on a journey to learn actuarial models without a rigorous understanding of such distributions. This poses a problem as even simple actuarial models make use of them.

This note is meant to bridge this gap by providing a self-contained monograph, peppered with problems, which deals with such distributions with mathematical rigor. The intent has been to make it easily accessible for self-study

This is the first version of a work in progress. Any comments will be appreciated - Please e-mail them to shyamal-kumar@uiowa.edu

Shyamal Kumar
30-Aug-2004

2 Distribution Functions

In this chapter we will first define and then study the properties of a distribution function.

Definition 1 A function $F(\cdot)$ is a distribution function if it satisfies the following three requirements:

a. $F(\cdot)$ is non-decreasing
b. $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$
c. $F(\cdot)$ is right continuous

To understand the definition, we recall that if a random variable, say X, is said to have $F(\cdot)$ as its distribution function, then

$$\Pr\left(X \in (-\infty, x]\right) = \Pr\left(X \leqslant x\right) = F(x), \quad \forall x \in \mathbb{R}$$

All the properties of a distribution function can now be derived from corresponding properties of a probability (measure). Since derivation will require a mathematical maturity beyond what is assumed and in any case will still lack motivation, we in the following motivate their presence.

Note that if $x \leqslant y$, then $(-\infty, x] \subseteq (-\infty, y]$. Hence, one would expect that

$$\Pr\left(X \in (-\infty, x]\right) \leqslant \Pr\left(X \in (-\infty, y]\right), \quad \forall x \leqslant y,$$

which is the monotonicity requirement.

Observing that

$$(-\infty, x] \to \mathbb{R}, \quad x \to \infty \quad \text{and} \quad (-\infty, x] \to \Phi, \quad x \to -\infty$$

and $\Pr(\mathbb{R}) = 1 = 1 - \Pr(\Phi)$, where Φ denotes the empty set, it would seem only natural to require

$$\Pr\left(X \in (-\infty, x]\right) \to 1, \quad \text{as } x \to \infty \quad \text{and} \quad \Pr\left(X \in (-\infty, x]\right) \to 0, \quad \text{as } x \to -\infty$$

This, in other words, is precisely $\lim_{x\to\infty} F(x) = 1 = 1 - \lim_{x\to-\infty} F(x)$.

The requirement for right continuity follows from the fact that

$$(-\infty, y] \downarrow (-\infty, x], \quad y \downarrow x.$$

As,

$$(-\infty, x - \frac{1}{n}] \to (-\infty, x), \quad n \to \infty$$

one does not require that distribution functions be left continuous. Nevertheless, $F(\cdot)$ has a left limit everywhere as result of it being non-decreasing. From above it will then be natural to expect that the left limit will satisfy

$$F(x-) \stackrel{\text{def}}{=} \lim_{\delta \downarrow 0} F(x-\delta) = \Pr\left(X \in (-\infty, x)\right), \quad \forall x.$$

This gives us the following lemma.

Lemma 1 If X is a random variable with $F(\cdot)$ as its distribution function, then

$$F(x) - F(x-) = \Pr\left(X = x\right), \quad \forall x.$$

Proof From the above, note that

$$F(x) - F(x-) = \Pr\left(X \in (-\infty, x]\right) - \Pr\left(X \in (-\infty, x)\right) = \Pr\left(X = x\right), \quad \forall x.$$

■

Definition 2 For a distribution function, say $F(\cdot)$, by $\mathcal{D}_F$ we shall denote its set of discontinuities. That is,

$$\mathcal{D}_F = \{x | F(x) - F(x-) > 0\}.$$

It can be shown that $\mathcal{D}_F$ is countable, and hence the following sum,

$$\sum_{x \in \mathbb{R}} F(x) - F(x-),$$

makes sense.

Problem 1** Show that $\mathcal{D}_F$ is at the most a countably infinite set.

Moreover, since $F(x) - F(x-)$ represents the height of the jump at x, the above sum represents the sum of the heights of all the jump discontinuities. We end this section with some problems.

Problem 2 The survival function of a random variable is defined as one minus its distribution function. Show that the survival function is a non-increasing function which is a also right continuous.

Problem 3 This problem explains why for random number generation, computers need to only be able to generate uniform random variables. Let F be a distribution function.

a. Define $F^{-1}(u) = \inf\{z | F(z) \geqslant u\}$, where inf stands for infimum. If you have not been exposed to infimum, then assume that for the problem it is the same as the minimum. And if you have been exposed, prove that the infimum is attained here for extra credit. Prove that

$$\{u | F^{-1}(u) \leqslant x\} = (-\infty, F(x)], \ \forall u \in (0, 1)$$

Hint:

$$\begin{aligned} u \leqslant F(x) &\Rightarrow x \in \{z | F(z) \geqslant u\} \\ &\Rightarrow x \geqslant \inf\{z | F(z) \geqslant u\} \\ &\Longleftrightarrow x \geqslant F^{-1}(u),\ \forall u \in (0,1) \end{aligned}$$

b. Let U be a $U(0,1)$ distributed random variable. Show that $F^{-1}(U)$ has F as it's distribution function.

c. Why did we have to define the *inverse* of F?

Problem 4 In the following assume the existence of all required moments.

a. It is easy to see that

$$\mathrm{Var}(X) = \min_{a \in \mathbb{R}} \mathbb{E}(X - a)^2$$

   i. Prove using Calculus
   ii. Prove without using Calculus

b. Let X and Y be two random variables.
   i. Show that $X - \mathbb{E}(X|Y)$ is uncorrelated with $\mathbb{E}(X|Y)$
   ii. Hence or otherwise show that

$$\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}(X|Y))^2 + \mathbb{E}(\mathbb{E}(X|Y) - \mathbb{E}(X))^2$$
$$\mathrm{Var}(X) = \mathbb{E}(\mathrm{Var}(X|Y)) + \mathrm{Var}(\mathbb{E}(X|Y))$$

   iii. Can you explain it as a **Pythagoras Theorem** - This part is for extra credit.

Problem 5 For the following functions, check if they satisfy the required properties of a distribution function.

A.

$$F_1(x) = \begin{cases} 0 & x < 1; \\ \frac{i(i+1)}{12} & i \le x < i+1, \quad i = 1, 2, 3; \\ 1 & \text{otherwise.} \end{cases}$$

B.

$$F_2(x) = \begin{cases} 0 & x < 1; \\ \frac{2ix - i(i+1)}{12} & i \le x < i+1, \quad i = 1, 2, 3; \\ 1 & \text{otherwise.} \end{cases}$$

C.

$$F_3(x) = \begin{cases} 0 & x < 1; \\ \frac{2ix - i(i+1)+6}{24} & i \le x < i+1, \quad i = 1,2,3; \\ 1 & \text{otherwise.} \end{cases}$$

Problem 6 Let X be a random variable which takes two values a and b with probabilities p and $q \stackrel{\text{def}}{=} 1 - p$.

a. Show that X is linearly related to a Bernoulli random variable
b. Using the above or otherwise find the mean and variance of X

3 Classification of Random Variables

In this chapter we shall identify two *two* classes of random variables - the random variables which to do not fit into these pure classes will be the third class, the class of mixed type.

Definition 3 A random variable will be called a continuous type random variable if its distribution function is continuous. In other words,

$$\sum_{x \in \mathbb{R}} [F(x) - F(x-)] = 0.$$

The above definition, in particular implies that the probability that a continuous random variable takes any particular value is zero and this in fact by itself is another definition of a continuous random variable. Also, another way to put is is that a continuous random variable has a distribution function with no jumps. On the other extreme are the discrete type random variables, defined below, which are *all about jumps*.

Definition 4 A random variable will be called a discrete type random variable random variable if its distribution function is such that,

$$\sum_{x \in \mathbb{R}} [F(x) - F(x-)] = 1.$$

From the above definitions, it should be clear that the random variables remaining are such that

$$0 < \sum_{x \in \mathbb{R}} [F(x) - F(x-)] < 1.$$

Such distributions, we term as of the mixed type.

Definition 5 A random variable will be called a mixed type random variable if its distribution function is such that,

$$0 < \sum_{x \in \mathbb{R}} [F(x) - F(x-)] < 1.$$

Problem 7 Prove that a random variable has to be one of continuous, discrete or mixed type.

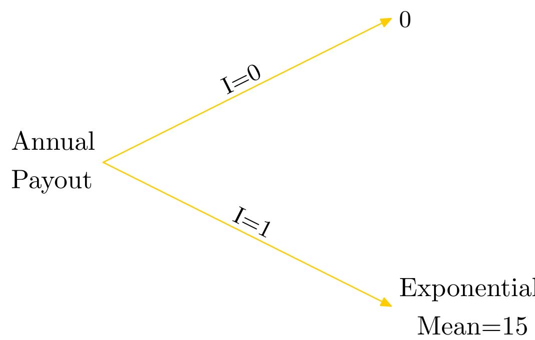
The practical interest in discrete and continuous random must be clear from elementary courses on Probability. The following example shows that mixed type random variables too arise in practice.

Example 1 A company sells toasters along with a warranty for one year against all repairs. It has been estimated that the probability of no repairs in the first year is 95%. In the case of repairs it has been found that the total payout in a year is distributed as a exponential with mean of $15.

Write down the distribution function of the payout under the warranty and classify the same. Also, find its expectation and variance.

Solution

We introduce a Bernoulli random variable, say I, with parameter 0.05 which will indicate the occurrence of a breakdown. Using this indicator random variable, the model can be depicted as done in the figure on the right. In words, conditional on the event $\{I = 0\}$ the annual payout is zero, and conditional on the event $\{I = 1\}$ the annual payout is distributed as an exponential with expectation of 15.



Let the annual payout be denoted by X. By Bayes theorem we have,

$$\Pr(X \leqslant x) = \Pr(I = 0) \cdot \Pr(X \leqslant x | I = 0) + \Pr(I = 1) \cdot \Pr(X \leqslant x | I = 1).$$

Moreover, from the description of the model we have

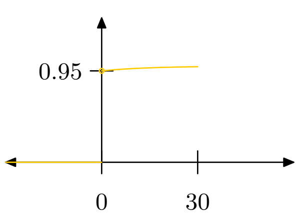
$$\Pr(X \leqslant x | I = 0) = \begin{cases} 0 & x < 0; \\ 1 & x \geqslant 0; \end{cases},$$

and

$$\Pr(X \leqslant x | I = 1) = \begin{cases} 0 & x < 0; \\ 1 - \exp\left\{\frac{-x}{15}\right\} & x \geqslant 0; \end{cases}.$$

The latter is the distribution function of an exponential distribution with mean of 15 and the former the distribution function of the degenerate distribution at 0. Combining the above we have,

$$\Pr(X \leqslant x) = \begin{cases} 0 & x < 0; \\ 1 - 0.05 \cdot \exp\left\{\frac{-x}{15}\right\} & x \geqslant 0; \end{cases}.$$



The distribution function of the annual payout has only jump, that at 0 with a height of 0.95. Hence according to the definition above it is of the mixed type. The distribution function is graphed on the right. Even though neither the methodology to calculate the moments of discrete nor continuous random variables can be put to use on their own, as we shall see in the next section, both of them can be applied to separate components of a mixed type distribution. Below we use an alternative approach.

Using the indicator random variable described above, we can write down

$$\mathbb{E}(X|I=i) = \begin{cases} 0 & i=0; \\ \\ 15 & i=1; \end{cases},$$

and

$$\mathrm{Var}\,(X|I=i) = \begin{cases} 0 & i=0; \\ \\ 225 & i=1; \end{cases}.$$

Hence,

$$\mathbb{E}(X) = 0.95 \cdot 0 + 0.05 \cdot 15 = 0.75$$

and using the formula of **Problem 4 b. ii.**, we have

$$\mathrm{Var}\,(X) = (0.95 \cdot 0 + 0.05 \cdot 225) + 0.95 \cdot 0.05 \cdot (0-15)^2 = 21.9375.$$

Observe that we have used the result of **Problem 6** above. ■

Problem 8 Classify the distributions in **Problem 5**.

Problem 9 There is one other way to find the moments of X in **Example 1**.

a. Let B be an exponential random variable with mean 15 and **independent** of I. Show that X can be written as $\mathrm{I} \cdot B$.
b. Using the above representation for X, especially the independence between I and B, find $\mathbb{E}(X)$, $\mathbb{E}\left(X^2\right)$ and using them or otherwise $\mathrm{Var}\,(X)$.

4 Decomposition of Mixed Type Distributions

In the example of the last section we saw that a discrete distribution (in that case a degenerate at zero) and a continuous distribution (in that case an exponential) can be combined to give rise to a distribution of the mixed type. The main goal of this section is to show that this is also a definition of a mixed-type distribution.

Let F be a mixed type distribution and let $X \sim \mathrm{F}$. By definition, $\mathcal{D}_\mathrm{F}$ is non-empty and moreover $\Pr(X \in \mathcal{D}_\mathrm{F}) \in (0,1)$. This makes both $\mathcal{D}_\mathrm{F}$ and $\mathcal{D}_\mathrm{F}^c$ non-trivial. By Bayes theorem then, we have

$$\Pr(X \leqslant x) = \Pr(X \in \mathcal{D}_\mathrm{F}) \cdot \Pr(X \leqslant x | X \in \mathcal{D}_\mathrm{F}) + \Pr\left(X \in \mathcal{D}_\mathrm{F}^c\right) \cdot \Pr\left(X \leqslant x | X \in \mathcal{D}_\mathrm{F}^c\right).$$

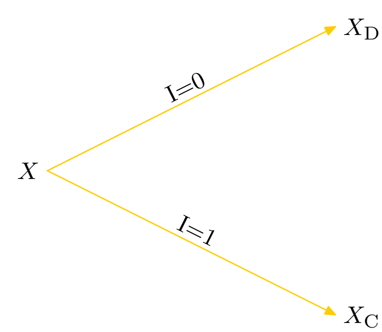
Let $\mathrm{F_D}(\cdot)$ and $\mathrm{F_C}(\cdot)$ be functions defined by

$$\mathrm{F_D}(x) = \Pr(X \leqslant x | X \in \mathcal{D}_\mathrm{F}) \quad \text{and} \quad \mathrm{F_C}(x) = \Pr\left(X \leqslant x | X \in \mathcal{D}_\mathrm{F}^c\right), \quad \forall x \in \mathbb{R}$$

and $\mathrm{p_D}$ and $\mathrm{p_C}$ be positive constants adding to one such that $\mathrm{p_D} = \Pr(X \in \mathcal{D}_\mathrm{F})$. Also, let X_D and X_C be random variables with $\mathrm{F_D}(\cdot)$ and $\mathrm{F_C}(\cdot)$ as their distribution functions, respectively. It can be shown that $\mathrm{F_D}(\cdot)$ and $\mathrm{F_C}(\cdot)$ are in fact distribution functions. But the results of the lemma is the reason for them being interesting.

Problem 10 Prove that $\mathrm{F_D}(\cdot)$ and $\mathrm{F_C}(\cdot)$ are distribution functions.

We shall first explain the results of the following lemma before stating and proving it. The lemma will say that with any mixed type random variable is associated a Bernoulli random variable with parameter $\mathrm{p_C}$ (coin toss), say I. If $\mathrm{I} = 1$ (coin lands head up) we say X is distributed as the continuous random variable X_C and on the other hand if $\mathrm{I} = 0$ (coin lands tail up) we say X is distributed as the discrete random variable X_D. Recall that in the example of the last section, with the newly introduced terminology, X_D happened to be a constant zero and X_C an exponential variate with mean of 15. One of the interesting results flowing from this description, actually used in the example referred to earlier, is encapsulated in the following problem.



Problem 11 If all the concerned moments exists, then

a.

$$\mathbb{E}\left(X^k\right) = \mathrm{p_D} \cdot \mathbb{E}\left(X_\mathrm{D}^k\right) + \mathrm{p_C} \cdot \mathbb{E}\left(X_\mathrm{C}^k\right), \quad \forall k.$$

b.

$$\mathrm{Var}(X) = (\mathrm{p_D} \cdot \mathrm{Var}(X_\mathrm{D}) + \mathrm{p_C} \cdot \mathrm{Var}(X_\mathrm{C})) + \mathrm{p_D} \cdot \mathrm{p_C} \cdot (\mathbb{E}(X_\mathrm{D}) - \mathbb{E}(X_\mathrm{C}))^2 .$$

Lemma 2 $\mathrm{F_D}(\cdot)$ and $\mathrm{F_C}(\cdot)$ defined above satisfy the following;

a. $\mathrm{F_D}(\cdot)$ is a discrete distribution function and moreover $\mathcal{D}_\mathrm{F}$ is the same as $\mathcal{D}_{\mathrm{F_D}}$.
b. $\mathrm{F_C}(\cdot)$ is a continuous distribution function.

Proof **Analysis of** $\mathrm{F_D}(\cdot)$**:** Observe that

$$\Pr(\{X_\mathrm{D} = x\}) = \Pr(X = x | X \in \mathcal{D}_\mathrm{F}) = \frac{\Pr(\{X = x\} \cap \{X \in \mathcal{D}_\mathrm{F}\})}{\Pr(X \in \mathcal{D}_\mathrm{F})},$$

which implies

$$\Pr(\{X_\mathrm{D} = x\}) = \begin{cases} 0 & x \in \mathcal{D}_\mathrm{F}^c \; ; \\ \frac{\Pr(\{X = x\})}{\Pr(X \in \mathcal{D}_\mathrm{F})} & x \in \mathcal{D}_\mathrm{F} \; ; \end{cases} .$$

The above can be simplified to

$$\Pr(\{X_\mathrm{D} = x\}) = \frac{\Pr(\{X = x\})}{\Pr(X \in \mathcal{D}_\mathrm{F})},$$

as $\Pr(\{X = x\}) = 0$ for all $x \in \mathcal{D}_\mathrm{F}^c$. Hence $\Pr(\{X_\mathrm{D} = x\}) > 0 \Leftrightarrow \Pr(\{X = x\}) > 0$ or in other words $\mathcal{D}_\mathrm{F} = \mathcal{D}_{\mathrm{F_D}}$. This in turn implies that

$$\sum [\mathrm{F_D}(x) - \mathrm{F_D}(x-)] = \sum_{x \in \mathcal{D}_{\mathrm{F_D}}} [\mathrm{F_D}(x) - \mathrm{F_D}(x-)] = \sum_{x \in \mathcal{D}_\mathrm{F}} [\Pr(X_\mathrm{D} = x)] = \sum_{x \in \mathcal{D}_\mathrm{F}} \left[\frac{\Pr(\{X = x\})}{\Pr(X \in \mathcal{D}_\mathrm{F})} \right] = 1,$$

or in other words, $\mathrm{F_D}(\cdot)$ is a discrete distribution function.

Analysis of $\mathrm{F_C}(\cdot)$**:** Observing that

$$\mathrm{F_C}(\cdot) = \frac{\mathrm{F}(\cdot) - \mathrm{p_D}\mathrm{F_D}(\cdot)}{\mathrm{p_C}},$$

and both $\mathrm{F}(\cdot)$ and $\mathrm{F_D}(\cdot)$ are left continuous on $\mathcal{D}_\mathrm{F}^c$, all we need to check is that $\mathrm{F_C}(\cdot)$ is left continuous on $\mathcal{D}_\mathrm{F}$. Using the above relation we have, in fact for all x,

$$\begin{aligned} \mathrm{F_C}(x) - \mathrm{F_C}(x-) &= \frac{(\mathrm{F}(x) - \mathrm{F}(x-)) - \mathrm{p_D}(\mathrm{F_D}(x) - \mathrm{F_D}(x-))}{\mathrm{p_C}} \\ &= \frac{\Pr(\{X = x\}) - \mathrm{p_D}\Pr(\{X_\mathrm{D} = x\})}{\mathrm{p_C}} \\ &= \frac{\Pr(\{X = x\}) - \Pr(X \in \mathcal{D}_\mathrm{F})\Pr(\{X_\mathrm{D} = x\})}{\mathrm{p_C}} \\ &= 0 \end{aligned}$$

which proves that $\mathrm{F_C}(\cdot)$ is continuous. ∎

Assumption: In this course we shall only deal with continuous distribution that permit a density. In the second version of this note there will be an elaboration of this point resulting in a star problem sequence on Cantor random variable, devil's stair case etc..

One of the implications of the last problem and the assumption is that the rules for the evaluation of the moments of discrete and continuous random variables can be applied to the respective parts of a mixed type random variable towards finding its moments. The following example will illustrate how to put this all into practice.

Example 2 Let $F(\cdot)$ be the distribution function given by

$$F(x) = \begin{cases} 0 & x < 1; \\ \frac{1}{2} + \frac{(x-1)}{4} & 1 \leqslant x < 2; \\ \frac{7}{8} + \frac{(x-2)}{8} & 2 \leqslant x < 3; \\ 1 & x \geqslant 3; \end{cases}.$$

a. Find the discrete and continuous components of F and also their weights (p_D and p_C).
b. Find the expectation and variance of F.

Solution We start by graphing the distribution function - this is done in the first figure below. The figure shows that there are two discontinuities, one each at 1 and 2.

The sum of the heights of the jumps is 0.652 (0.5 + 0.125) and this is p_D. Once we know p_D, the probability mass function of X_D is calculated by finding the **relative** height of the jumps, relative to the total height of the jumps. This leads to,

$$\Pr(\{X_D = x\}) = \begin{cases} \frac{0.5}{0.625} = 0.8 & x = 1; \\ \frac{0.125}{0.625} = 0.2 & x = 2; \end{cases}.$$

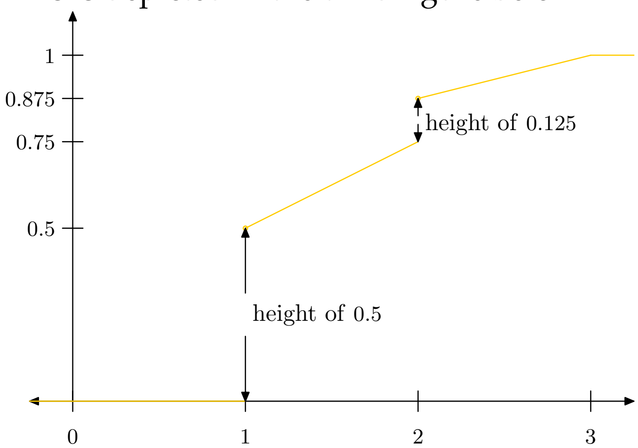
The distribution function corresponding to the above probability mass function is graphed as the second figure below. The distribution function can be written as

$$F_D(x) = \begin{cases} 0 & x < 1; \\ 0.8 & 1 \leqslant x < 2; \\ 1 & x \geqslant 2; \end{cases}.$$

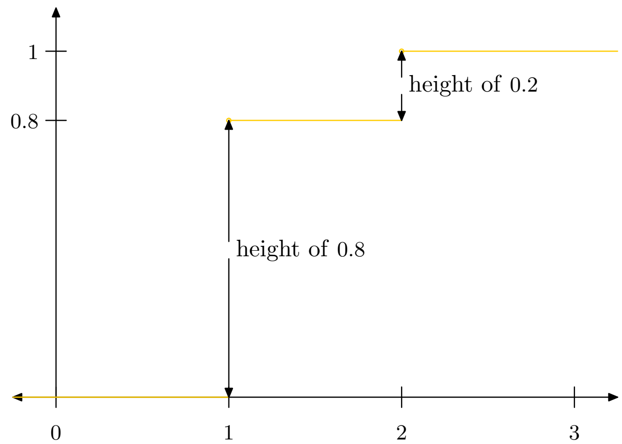
Once $F_D(\cdot)$ is known, by algebra one can write down $F_C(\cdot)$ - in this case it turns out to be

$$F_C(x) = \begin{cases} 0 & x < 1; \\ \frac{2(x-1)}{3} & 1 \leqslant x < 2; \\ \frac{(x-2)}{3} & 2 \leqslant x < 3; \\ 1 & x \geqslant 3; \end{cases}.$$

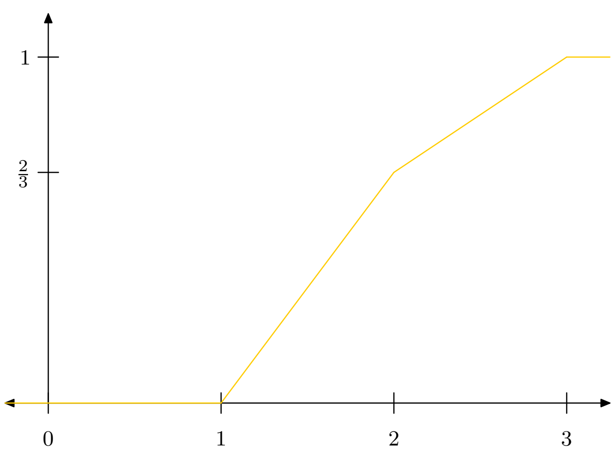
This is depicted in the third figure below.



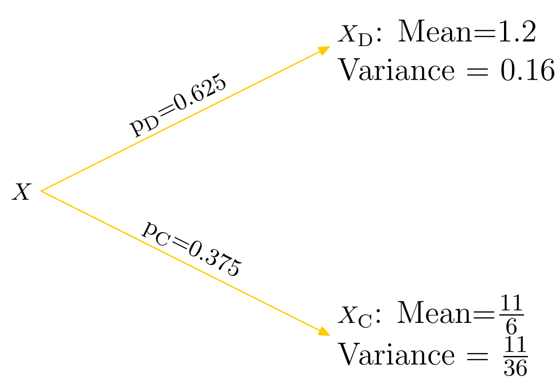
i: F(·) - Mixed Type



ii: Discrete Component



iii: Continuous Component



iv: The Weights & the Moments

The density of the continuous component can be found by differentiating the above, which gives rise to

$$f_C(x) = \begin{cases} 0 & x < 1; \\ \frac{2}{3} & 1 \leqslant x < 2; \\ \frac{1}{3} & 2 \leqslant x < 3; \\ 0 & x \geqslant 3; \end{cases}.$$

Now with the distributions of X_D and X_C well specified, we can find their moments by using methods for discrete and continuous distributions respectively. This results in the fourth figure above. Now the mean and variance of F can be calculated as

$$\begin{aligned} E(X) &= p_D \cdot \mathbb{E}(X_D) + p_C \cdot \mathbb{E}(X_C) \\ &= \frac{5}{8} \cdot \frac{6}{5} + \frac{3}{8} \cdot \frac{11}{6} \\ &= \frac{23}{16} \end{aligned}$$

and

$$\begin{aligned} \mathrm{Var}\,(X) &= (p_D \cdot \mathrm{Var}\,(X_D) + p_C \cdot \mathrm{Var}\,(X_C)) + p_D \cdot p_C \cdot (\mathbb{E}(X_D) - \mathbb{E}(X_C))^2 \\ &= \left(\frac{5}{8} \cdot \frac{4}{25} + \frac{3}{8} \cdot \frac{11}{36}\right) + \frac{5}{8} \cdot \frac{3}{8} \cdot \left(\frac{6}{5} - \frac{11}{6}\right)^2 \\ &\simeq 0.31. \end{aligned}$$

■