Gary R. Skoog and James E. Ciecka. 2011. Probability Mass Functions in Forensic Economics. *Journal of Legal Economics* 18(1): pp. 95-107.

Probability Mass Functions in Forensic Economics

Gary R. Skoog and James E. Ciecka *

This note is about the concept of a probability mass function (pmf) and two specific pmfs that have become part of the forensic economics literature. This note has the following objectives. It explains the concept of a pmf within the framework of clear examples. It provides two examples of pmfs within the contexts most likely to be encountered by, and used by, forensic economists. It presents examples of pmfs that are rich enough to illustrate essential ideas but do not get bogged down in unnecessary details. It aims to clearly state what a pmf can and cannot do for forensic economists.

We begin with the following definition of a pmf:

> *A pmf is a statement that relates discrete events to the probabilities associated with those events occurring*.

Probabilities must be greater than or equal to zero and sum to 1.00. The word "mass" is used to indicate that probabilities are concentrated on discrete events.[1]

Consider the following examples:

Example 1. Suppose there are 100 balls in a drum; 90 balls are marked with $5 and the remaining 10 balls are marked with $20. A ball is randomly drawn from the drum. The pmf consists of the following statement:

> *The probability is .90 that a draw results in a $5 ball, and the probability is .10 that a draw yields a ball with $20 on it.*

This pmf is illustrated in Figure 1 in the Figures and Tables section at the end of this note.

---

* Gary R. Skoog, Department of Economics, DePaul University, Chicago IL; and Legal Econometrics, Glenview, IL. James E. Ciecka, Department of Economics, DePaul University, Chicago IL. Phone: 312 362-8831, E-mail: jciecka@depaul.edu. Our thanks to Thomas R. Ireland for suggesting that we write this note. We thank Lane Hudgins, Thomas R. Ireland, Michael L. Nieswiadomy, and Marc Weinstein for helpful comments.

The expected value of a random draw is equal to the sum of each possible outcome of a draw multiplied by its probability of occurring. In this example, Expected Value = \$5(.9) + \$20(.1) = \$6.50. Note that the expected value, as shown in this example, does not have to be equal to any of the possible outcomes that can occur.

Example 2. Consider a group of 100,000 people at exact age 30. (The word "exact" means that they have just had their 30th birthday.) Some people in this group will die before reaching their 31st birthday. Deaths occur uniformly over the year; that is, some will die very shortly after their 30th birthday, some die a few months later, and some die close to their 31st birthday. With deaths occurring uniformly throughout the year, it can be said that, on average, deaths occur at the midpoint between exact ages 30 and 31; and in general deaths occur at the midpoint between any other adjacent exact ages. This example assumes that the probability of a 30 year old dying between ages 30 and 31 is .20; for those alive at age 31, the probability of dying between ages 31 and 32 is .60; and everyone alive at age 32 dies between ages 32 and 33. With the midpoint assumption, deaths occur at ages 30.5, 31.5 and 32.5.

Of the 100,000 people initially alive at exact age 30, 20,000 people (.20 of the original 100,000) die at age 30.5. That is, *the probability is .20 for living exactly .5 years beyond age 30*. There remain 80,000 survivors who live past age 30.5 and 60 percent of these, or 48,000, die at age 31.5. These 48,000 who die at age 31.5 have lived 1.5 years beyond age 30; and they represent .48 of the original population of 100,000 people. That is, *the probability is .48 for living exactly 1.5 years beyond age 30*. There are 32,000 survivors after age 31.5 because 20,000 have died at age 30.5 and another 48,000 have died at age 31.5. All 32,000 survivors, or .32 of the original 100,000, die at age 32.5. Those dying at age 32.5 lived 2.5 years beyond age 30. That is, *the probability is .32 for living exactly 2.5 years beyond age 30.*

The italicized statements in the previous paragraph comprise the pmf for additional years of life. The statements are:

> *The probability is .20 for living exactly .5 years beyond age 30, the probability is .48 for living exactly 1.5 years beyond age 30, and the probability is .32 for living exactly 2.5 years beyond age 30.*

All probability masses for additional years of life sum to 1.00 = .20 + .48 +.32. The pmf is illustrated in Figure 2. The expected value of additional years of life, or life expectancy at exact age 30, is *LE* = .5(.20) + 1.5(.48) + 2.5(.32) = 1.62 years. The median and mode of additional years of life are 1.5 years. The standard deviation is $\sqrt{.20(.5-1.62)^2 + .48(1.5-1.62)^2 + .32(2.5-1.62)^2} =$ .71 years.

Example 3. A person who is active in the labor force can remain active in the future or become inactive; similarly, an inactive person can remain inactive or become active. In addition, an active or inactive person can die in the future. As in Example 2, transitions from one labor force state to another, or to the death state, are assumed to always occur at a midpoint between a person's birthdays. This example will consider people who are exactly age 30 and who live to at most age 32.5. All transition points between activity and inactivity and transitions to death occur at ages 30.5, 31.5, and 32.5. Assume the following probabilities:

(a) The active-to-active transition probability between ages 30 to 31 and between ages 31 to 32 is .70.

(b) The active-to-inactive transition probability between ages 30 to 31 and between ages 31 to 32 is .20.

(c) The inactive-to-active transition probability between ages 30 to 31 and between ages 31 to 32 is .50.

(d) The inactive-to-inactive transition probability between ages 30 to 31 and between ages 31 to 32 is .40.

(e) The probability of a transition to death between ages 30 to 31 and between ages 31 to 32 is .10 regardless of whether a person is active or inactive before dying. Once a person reaches age 32.5 death occurs with certainty because no one lives to age 33. Therefore, the probability of death at age 32.5 is 1.00.

Assumptions (a) through (d) create seven possible labor force paths.[2] Each path occurs with a certain probability, and each path has a specific number of years of labor force activity associated with it. We consider each path, for an initially active person, separately below:

Path 1. A person is active at age 30; the person passes over a transition point at age 30.5 but stays active; the person passes over a transition point at age 31.5 but stays active; the person reaches a transition point at age 32.5 and dies. This path results in 2.5 years of labor force activity, and it occurs with probability (.7)(.7)(1.00) = .49.

Path 2. A person is active at age 30; the person passes over a transition point at age 30.5 but stays active; the person becomes inactive at the age 31.5 transition point; the person reaches a transition point at age 32.5 and dies. This path results in 1.5 years of labor force activity, and it occurs with probability (.7)(.2)(1.00) = .14.

Path 3. A person is active at age 30; the person becomes inactive at the age 30.5 transition point; the person becomes active at the age 31.5 transition point; the person reaches a transition point at age 32.5 and dies. This path results in 1.5 years of labor force activity, and it occurs with probability (.2)(.5)(1.00) = .10.

Path 4. A person is active at age 30; the person becomes inactive at the age 30.5 transition point; the person passes over a transition point at age 31.5 but stays inactive; the person reaches a transition point at age 32.5 and dies. This path results in .5 years of labor force activity, and it occurs with probability (.2)(.4)(1.00) = .08.

Path 5. A person is active at age 30; the person passes over a transition point at age 30.5 but stays active; the person reaches a transition point at age 31.5 and dies. This path results in 1.5 years of labor force activity, and it occurs with probability (.7)(.1) = .07.

Path 6. A person is active at age 30; the person becomes inactive at the age 30.5 transition point; the person reaches a transition point at age 31.5 and dies. This path results in .5 years of labor force activity, and it occurs with probability (.2)(.1) = .02.

Path 7. A person is active at age 30; the person reaches a transition point at age 30.5 and dies. This path results in .5 years of labor force activity, and it occurs with probability .10.

The information for all seven paths is summarized in Table 1. Note that Paths 4, 6, and 7 occur with probabilities .08, .02, and .10, but they all produce .5 years of labor force activity. Therefore, .5 years of activity occurs with combined probability of .08 + .02 + .10 = .20. Paths 2, 3, and 5 occur with probabilities .14, .10, and .07, but they all produce 1.5 years of labor force activity. Therefore, 1.5 years of activity occurs with combined probability of .14 + .10 + .07 = .31. Only Path 1 results in 2.5 years of activity which occurs with probability .49. Now it is possible to state the pmf for future years of labor market activity for a person initially active at exact age 30:

> *The probability is .20 for .5 years of additional labor market activity beyond age 30; the probability is .31 for 1.5 years of additional labor market activity beyond age 30; and the probability is .49 for 2.5 years of additional labor market activity beyond age 30.*

This pmf is illustrated in Figure 3.

Worklife expectancy (*WLE*) is the expected value of additional years of activity: *WLE* = .5(.20) + 1.5(.31) + 2.5(.49) = 1.79 years. In addition to *WLE*, the pmf allows us to compute other parameters of additional years in the labor market. For example, the median is 1.50 years of activity and the mode is 2.5 years. The standard deviation of years of activity is $\sqrt{.20(.5-1.79)^2+.31(1.5-1.79)^2+.49(2.5-1.79)^2}=$ .78 years. [See Skoog and Ciecka (2001a, 2001b, and 2002) and Skoog, Ciecka, and Krueger (2011) for applications of this type of pmf.]

Path 2 and Path 3 are similar in the sense that they both produce 1.5 additional years of activity. However, their time configurations differ. In Path 2, activity commences at age 30 and runs uninterrupted for 1.5 years until age 31.5. Path 3 begins with .5 years of activity between ages 30 and 30.5, an interruption occurs between ages 30.5 and 31.5 during which there is no activity, and the activity resumes between ages 31.5 and 32.5. The present value of a dollar of loss is greater along Path 2 than Path 3 because more activity occurs earlier along Path 2. This exemplifies the following point: The pmf, or *WLE* which is calculated with the pmf, cannot be used to compute present value. Present value calculations require the exact decomposition of *WLE* as shown by Skoog (2001) and Skoog and Ciecka (2006 and 2009) or with close approximations (Skoog and Ciecka, 2006).

Although Example 2 and Example 3 are based on a very short age interval and high mortality, they capture important properties of pmfs. Probability mass functions enable us to compute life expectancies and worklife expectancies; but, if that were all they could do for forensic economists, pmfs would be of little value because these expectations have been computed for many years without using pmfs. However, other characteristics like the median and various percentile points cannot be correctly calculated for years of activity without a pmf; nor can the mode, standard deviation, skewness, and kurtosis be calculated. That is, pmfs enable forensic economists to move beyond worklife expectancies and thereby understand much more about the characteristics of years of activity. There are limitations. Example 3 shows that more than one activity/inactivity/death path can produce the same number of years of future labor activity. Not all equal-years-of-activity paths yield the same present value of a dollar in lost earnings because activity can occur at different times in the future as shown in Path 2 and Path 3.

This note ends with an illustration in Figure 4 of a pmf function for 35-year-old active men with a BA degree from Skoog, Ciecka, and Krueger (2011). *WLE* is 27.52 years, the median is 28.50 years, and the mode is 29.50 years. The standard deviation, skewness, and kurtosis are 7.34 years, -.63 years, and 4.06 years, respectively. The 10^{th}, 25^{th}, 75^{th}, and 90^{th} percentile points are 18.50 years, 23.50 years, 32.50 years, and 35.50 years, respectively. Although the pmf in Figure 4 looks roughly normal, the mean (*WLE*) is less than the median which is less than the mode. Years of future activity vary substantially (the standard deviation is 7.34 years), the slight negative skewness indicates more probability mass in the left tail, and the kurtosis of 4.06 indicates that there is more probability mass in the peak region and the tails than a normally distributed variable.[3] There are two main differences between Figure 4 and Figure 3: (1) Figure 4 is based on actual transition and mortality probabilities while Figure 3 is based on hypothetical data and (2) Figure 4 captures $2^{111-35} - 1 = 7.55 \times 10^{22}$ labor force paths and Figure 3 is based on $2^{33-30} - 1 = 7$ paths.[4] These differences notwithstanding, Figure 3 and Figure 4 rest on the same fundamental ideas.

Figures and Tables

Figure 1. PMF for Random Draw of a Ball

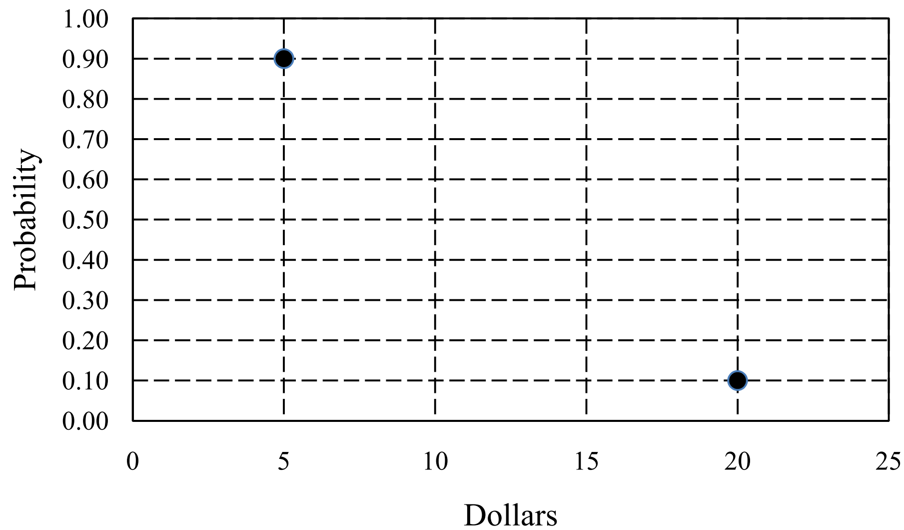

Figure 2. PMF for Additional Years of Life

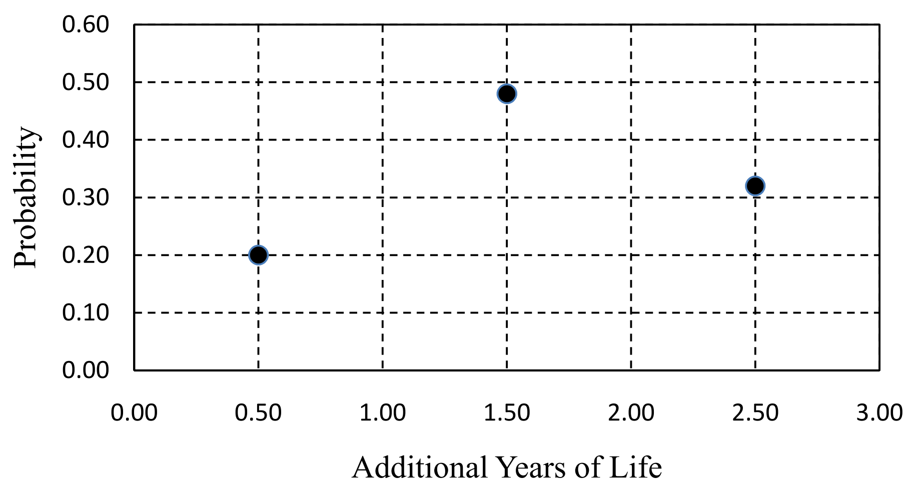

Figure 3. PMF for Additional Years of Labor Market Activity

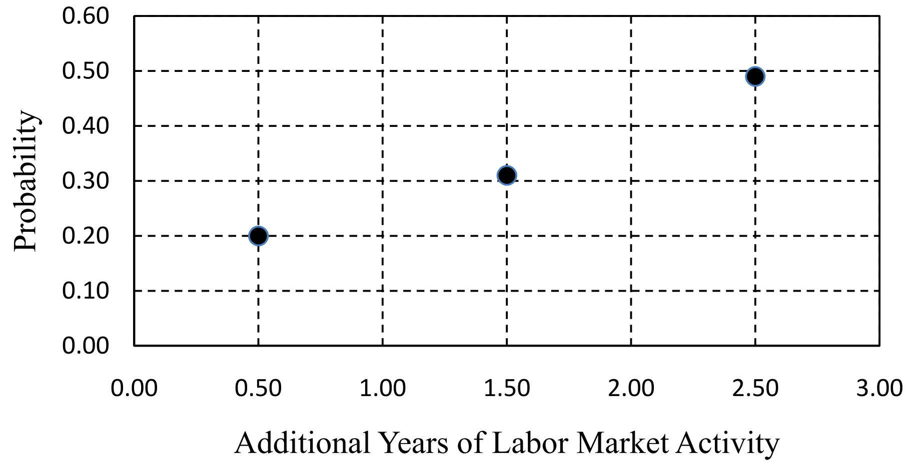

Figure 4. PMF for Active Men Age 35 with BA Degrees

Table 1. Labor Force Paths and Their Associated Probabilities of Occurrence and Years of Labor Force Activity

Paths	Probability	Years of Activity
Path 1	.49	2.5
Path 2	.14	1.5
Path 3	.10	1.5
Path 4	.08	.5
Path 5	.07	1.5
Path 6	.02	.5
Path 7	.10	.5
Total	1.00	

Endnotes

1. There is some intuition behind the choice of the word "mass" in the term "probability mass function." To develop this intuition, it is useful to think about a continuous probability function $f(x)$ – a normally distributed random variable would be a good example. Interpret $f(x)$ (i.e., the height of the continuous probability function) as probability per unit of x. Note that in the continuous case, $f(x)$ is not the probability of x occurring, but rather probability per unit of x. To compute probability, one must multiply $f(x)$ by a number of units of x. For example, suppose IQ scores follow a continuous probability function and let x=110 and assume $f(x) = f(110) = .05$ where the latter number is probability per unit of x. To determine the probability of a person's IQ score being between 109 and 111 (i.e., a range of $\Delta x = 2$ units centered around x=110) compute $f(x)\Delta x = f(110)(2) = .05(2) = .10$. It can be said that the probability of having an IQ between 109 and 111 is .10. (This is only an approximate answer; the exact probability would be computed by integrating $f(x)$ over the range from 109 to 111.) The fundamental idea is that probability depends on the range in x from 109 to 111 and how "dense" or concentrated probability per unit of x happens to be in the range of IQ scores from 109 to 111. This is the motivation for calling $f(x)$ a probability "density" function in the case of a continuous random variable. So, probability is $f(x)\Delta x$; or, in words, probability is density multiplied by a change in a variable. Finally, observe that the word "mass" is defined as the product of density and a change in a variable – i.e., in the context this note is considering, the word "mass" denotes probability. In the case of a discrete random variable, one does not have to multiply density by a change in x in order to get probability (i.e., mass); units measured on the vertical axis are already probabilities. The probability function in a discrete situation is called a "mass" function because its units are probabilities and "mass" conveys the idea of probability.

2. For simplicity this note assumes that transitions from active to active occur with the same probability between ages 30 to 31 and between ages 31 to 32. This note makes similar constancy assumptions for active to inactive, inactive to active, and inactive to inactive transitions between ages 30 to 31 and 31 to 32. Of course, transition probabilities need not be constant.

3. Picture a normal probability density function. This note refers to the region in the middle where the function is the highest as the "peak" region and refers to the regions in the edges where the function is the lowest as the "tails." The parts between either tail and the peak region are referred to as the "shoulders" of the function. It turns out that the kurtosis

of a normal random variable is always exactly 3.00 regardless of how small or large its standard deviation may be. The normal probability density function is the standard for thinking about kurtosis. A random variable is said to be "leptokurtic" if its kurtosis exceeds 3.00. Such a variable tends to have a higher "peak," thicker and longer "tails," and lower "shoulders" than a normal variable. That is, there is more probability material in the "peak" and "tail" regions and less probability material in the "shoulders" than a normal variable. A "platykurtic" random variable has the opposite appearance. It has a lower "peak," shorter "tails," and higher "shoulders" than a normal random variable. The kurtosis measure is less than 3.00 for a platykurtic variable.

Kurtosis can be important. For example, before the recent financial crisis, the world's largest financial firms used the normal probability density function in their risk management models. These models produced predictions of extreme events (i.e., events far into the tails) as very, very unlikely. However, the probability of extreme events like huge negative returns on investments is larger if a variable is leptokurtic. Since the so called "Black Swan" event is more probable if returns are leptokurtic, a more conservative investment strategy should have been pursued than what appeared appropriate using the normality assumption.

Characteristics of random variables are computed in order to understand their behavior. The most basic characteristics are mean, median, mode, standard deviation, skewness, and kurtosis. In regard to Figure 4, the pmf for active 35-year-old men with BA degrees is leptokurtic because its kurtosis measure is 4.06 > 3.00. This means, relative to normality, that there is more probability for years of activity in the peak region (where the worklife expectancy, median, and mode are located) and more probability in its tails. Therefore, this pmf tends to be higher peaked and thicker tailed than a normally distributed variable. This is one, among several, characteristics of years of activity for active 35-year-old men with BA degrees.

4. The number of paths is equal to 1 subtracted from 2 raised to the power calculated as the youngest exact age at which everyone has died less current exact age. The youngest exact age at which everyone has died is 111 for the pmf in Figure 4 (Skoog, Ciecka, and Krueger, 2011). Current age is 35, so the number of paths is $2^{111-35}-1=7.55\text{x}10^{22}$. In Example 3, the youngest exact age at which everyone has died is 33, and current exact age is 30. Therefore, the number of paths is $2^{33-30}-1=7$. There are a huge number of paths even if we were to assume that everyone were dead at exact age as young as 80. Consider, for example a person exactly age 30, the number of paths would be $2^{80-30}-1=1.12\text{x}10^{15}$. There is no practical way to trace such a large number of paths. However, Skoog and Ciecka (2002) found recursions which compute exact probabilities of

years of activity as if each path were actually searched, its probability calculated, and assigned to the number of years of activity in the path.

References

Skoog, Gary, R. 2001. Worklife Expectancy: Theoretical Results. presented at Allied Social Science Association meeting, Atlanta.

Skoog, Gary R. and James E. Ciecka. 2001a. The Markov (Increment-Decrement) Model of Labor Force Activity: New Results Beyond Worklife Expectancies. *Journal of Legal Economics*, 11(1), 1-21.

________. 2001b. The Markov (Increment-Decrement) Model of Labor Force Activity: Extended Tables of Central Tendency, Variation, and Probability Intervals. *Journal of Legal Economics*, 11(1), 23-87.

________. 2002. Probability Mass Functions for Labor Force Activity Induced by the Markov (Increment-Decrement) Model of Labor Force Activity. *Economics Letters*, 77(3), 425-431.

______. 2006 Allocation of Worklife Expectancy and the Analysis of Front and Uniform Loading with Nomograms. *Journal of Forensic Economics*, 19(3), 261-296.

______. 2009. Present Value Functions and Recursions. *Journal of Forensic Economics*, 21(1), 63-98.

________ and Kurt Krueger. 2011. The Markov Process Model of Labor Force Activity: Extended Tables of Central Tendency, Shape, Percentile Points, and Bootstrap Standard Errors. *Journal of Forensic Economics*, 22(2), 165-229.