

General random variables

Identifying the PDF

A recent college graduate is moving to Houston, Texas to take a new job, and is looking to purchase a home. Since Houston comprises a relatively large metropolitan area of about 7 million people, there are a lot of homes from which to choose. When searching for properties on the real estate websites, it is possible to select the price range of housing in which one is most interested. Suppose the potential buyer specifies a price range of \$200,000 to \$250,000, and the result of the search returns thousands of homes with prices distributed uniformly throughout that range. Identify $E[X]$ and σ of the probability density function associated with this random variable.¹

Solution:

$$\begin{aligned} E[X] &= \frac{a+b}{2} \\ &= \frac{200,000 + 250,000}{2} \\ &= \frac{450,000}{2} \\ &= 225,000 \end{aligned}$$

$$\begin{aligned} Var[X] &= \frac{(b-a)^2}{12} \\ &= \frac{(250,000 - 200,000)^2}{12} \\ &= \frac{50,000^2}{12} \\ &= 2.083 \times 10^8 \end{aligned}$$

$$\begin{aligned} Var[X] &= \sigma^2 \\ sd(X) &= \sigma \\ &= \sqrt{2.083 \times 10^8} \\ &= 1.443 \times 10^4 \end{aligned}$$

¹Stinerock 6.1

Calculating ideal points

Suppose Bob is a voter living in the country of Freedonia, and suppose that in Freedonia, all sets of public policies can be thought of as representing points on a single axis (e.g. a line running from more liberal to more conservative). Bob has a certain set of public policies that he wants to see enacted. This is represented by point v , which we will call Bob's **ideal point**. The utility, or happiness, that Bob receives from a set of policies at point l is $U(l) = -(l - v)^2$. In other words, Bob is happiest if the policies enacted are the ones at his ideal point, and he gets less and less happy as policies get farther away from his ideal point. When he votes, Bob will pick the candidate whose policies will make him happiest. However, Bob does not know exactly what policies each candidate will enact if elected – he has some guesses, but he can't be certain. Each candidate's future policies can therefore be represented by a continuous random variable L with expected value μ_l and variance $Var(L)$.²

- a. Express $E(U(L))$ as a function of μ_l , $Var(L)$, and v . Why might we say that Bob is **risk averse** – that is, that Bob gets less happy as outcomes get more uncertain?

Solution: In order to express $E(U(L))$ as a function of μ_l , $Var(L)$, and v , we need to find a form of the equation that contains those variables and no others. We know that

$$\begin{aligned}
 E(U(L)) &= E(-(L - v)^2) \\
 &= E(-(L^2 - 2vL + v^2)) \\
 &= E(-L^2 + 2vL - v^2) \\
 &= \int_{-\infty}^{\infty} (-l^2 + 2vl - v^2)f(l) dl \\
 &= \int_{-\infty}^{\infty} -l^2 f(l) dl + \int_{-\infty}^{\infty} 2vl f(l) dl + \int_{-\infty}^{\infty} -v^2 f(l) dl \\
 &= -\int_{-\infty}^{\infty} l^2 f(l) dl + 2v \int_{-\infty}^{\infty} l f(l) dl - v^2 \int_{-\infty}^{\infty} f(l) dl \\
 &= -E(L^2) + 2vE(L) - v^2 \\
 &= -E(L^2) + 2v\mu_l - v^2
 \end{aligned}$$

Notice, however, that $E(L^2) = Var(L) + E(L)^2$, which can be determined by adding $E(L)^2$ to both sides of the definition of $Var(L)$. Therefore

$$\begin{aligned}
 E(U(L)) &= -E(L^2) + 2v\mu_l - v^2 \\
 &= -Var(L) - \mu_l^2 + 2v\mu_l - v^2
 \end{aligned}$$

Bob can be considered risk averse because his expected utility is inversely related to the variance of L ; as the variance of potential policy outcomes increases – or, in other words, as the candidate becomes a riskier bet – the less utility he receives. This problem illustrates one way in which the concepts we've been talking about, including random variables, expected value, and variance, can be applied to formal or game theoretic analysis.

²Grimmer HW11.2

- a. Suppose Bob is deciding whether to vote for one of two candidates: Dwight Schrute or Leslie Knope. Suppose Bob's ideal point is at 1, Schrute's policies can be represented by a continuous random variable L_S with expected value at 1 and variance equal to 6, and Knope's policies can be represented by a continuous random variable L_K with expected value at 3 and variance equal to 1. Which candidate would Bob vote for and why? What (perhaps surprising) effect of risk aversion on voting behavior does this example demonstrate?

Solution: To see who Bob will choose, we first calculate the utility he derives from each of the potential recipients of his vote.

$$\begin{aligned} E(U(L_T)) &= -\text{Var}(L_T) - E(L_T)^2 + 2vE(L_T) - v^2 \\ &= -6 - (1^2) + 2(1)(1) - (1^2) \\ &= -6 \end{aligned}$$

$$\begin{aligned} E(U(L_W)) &= -\text{Var}(L_W) - E(L_W)^2 + 2vE(L_W) - v^2 \\ &= -1 - (3^2) + 2(1)(3) - (1^2) \\ &= -1 - 9 + 6 - 1 \\ &= -5 \end{aligned}$$

Notice that $E(U(L_W)) > E(U(L_T))$. Therefore Bob will vote for Joe Biden. Although his ideal point is closer to the expected value of Trump's policies, the greatly reduced uncertainty surrounding Biden's potential policies is sufficient to convince Bob to vote for her instead. More generally, risk aversion can in some cases induce individuals to vote for candidates whose mean potential policies are not the closest to those individuals' ideal points.

Parliamentary elections

After an election in a parliamentary system, a government (consisting of a prime minister and a cabinet) is formed by gathering the support of a majority of newly elected members of parliament. Typically a government is allowed to remain in power for a certain number of years before new elections must be called. However, elections can be held earlier if the Parliament passes a vote of no confidence or the prime minister decides to dissolve the government. Suppose we are studying Country Z (which uses a parliamentary system) and we are interested in the duration of governments. In Country Z, governments must call elections at least every 5 years, but they could be called sooner if there is a vote of no confidence or the prime minister dissolves the government. Let the continuous random variable X denote the amount of time (measured in years) between the last election and the calling of the next election. X has support on all real numbers between 0 and 5. Suppose we know that X has the probability density function

$$f(x) = \begin{cases} kx^3 & 0 < x < 5 \\ 0 & \text{otherwise} \end{cases}$$

where k is some constant.

a. Find k .

Solution: Recall that a valid PDF must sum up to 1. So we have to find a k such that the integral of the PDF equals 1.

$$\begin{aligned}\int_0^5 kx^3 dx &= 1 \\ k \int_0^5 x^3 dx &= 1 \\ k \left[\frac{x^4}{4} \right]_0^5 &= 1 \\ k \left[\frac{5^4}{4} - \frac{0^4}{4} \right] &= 1 \\ \frac{625k}{4} &= 1 \\ k &= \frac{4}{625}\end{aligned}$$

b. Find the CDF of X .

Solution:

$$\int_{-\infty}^x \frac{4x^3}{625} dx = \int_0^x \frac{4x^3}{625} dx = \frac{4}{625} \int_0^x x^3 dx = \frac{4}{625} \left[\frac{x^4}{4} \Big|_0^x \right] = \frac{x^4}{625}$$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x^4}{625} & \text{if } 0 < x < 5 \\ 1 & \text{if } x > 5 \end{cases}$$

Note that, since technically $-\infty < x < \infty$, you should define for CDF for values of x above and below the bounds are you primarily interested in.

a. Find $E(X)$ and $Var(X)$.

Solution:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x \cdot \frac{4}{625} \cdot x^3 dx \\ &= \int_0^5 x \cdot \frac{4}{625} \cdot x^3 dx \\ &= \frac{4}{625} \int_0^5 x^4 dx \\ &= \frac{4}{625} \left[\frac{x^5}{5} \Big|_0^5 \right] \\ &= \frac{4}{625} \cdot 625 \\ &= 4 \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 \cdot \frac{4}{625} \cdot x^3 dx \\ &= \int_0^5 x^2 \cdot \frac{4}{625} \cdot x^3 dx \\ &= \frac{4}{625} \int_0^5 x^5 dx \\ &= \frac{4}{625} \left[\frac{x^6}{6} \Big|_0^5 \right] \\ &= \frac{4}{625} \cdot \frac{15625}{6} \\ &= 25 \cdot \frac{2}{3} \\ &= \frac{50}{3} \end{aligned}$$

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{50}{3} - 4^2 = \frac{50}{3} - 16 = \frac{2}{3}$$

a. Find the median of X (the value of x at which $\Pr(X \leq x) = \frac{1}{2}$).

Solution:

$$\begin{aligned}
\int_{-\infty}^m \frac{4}{625} \cdot x^3 dx &= \frac{1}{2} \\
\int_0^m \frac{4}{625} \cdot x^3 dx &= \frac{1}{2} \\
\frac{4}{625} \int_0^m x^3 dx &= \frac{1}{2} \\
\frac{4}{625} \left[\frac{x^4}{4} \right]_0^m &= \frac{1}{2} \\
\frac{4}{625} \cdot \frac{m^4}{4} &= \frac{1}{2} \\
m^4 &= \frac{625}{2} \\
m &= \sqrt[4]{\frac{625}{2}} \\
&= \frac{5}{\sqrt[4]{2}} \approx 4.2045
\end{aligned}$$

- b. What is the probability that the government remains in power for exactly 3 years? Why?

Solution: Zero. If the random variable is continuous, the probability of any one particular outcome is zero.

- c. What is the probability that the government remains in power between 2 and 4 years?

Solution:

$$F(4) - F(2) = \frac{4^4}{625} - \frac{2^4}{625} = \frac{256 - 16}{625} = \frac{240}{625} = 0.384$$

The probability that the government remains in power between 2 and 4 years is 0.384.

d. What is the probability that the government remains in power for less than 1 year or more than 4 years?

Solution: This is the probability that the government survives between 0 and 1 year, $F(1)$, plus the probability that the government survives between 4 and 5 years, $F(5) - F(4) = 1 - F(4)$.

$$\begin{aligned} F(1) + [1 - F(4)] &= \frac{1^4}{625} + \left[1 - \frac{4^4}{625}\right] = \frac{1}{625} + \left[1 - \frac{256}{625}\right] \\ &= \frac{1}{625} + \frac{369}{625} = \frac{370}{625} = 0.592 \end{aligned}$$

The probability that the government remains in power for less than one year or more than 4 years is 0.592.

Calculating the CDF

Z is distributed according to the following PDF

$$f(z) = \begin{cases} \gamma \exp(-\gamma z) & 0 \leq z \\ 0 & \text{otherwise} \end{cases}$$

- a. What is $F(z)$, the CDF of this distribution?

Solution: Z is only defined for the interval when greater than 0, so to find the CDF we assess $P(Z \leq z)$.

$$\begin{aligned} F(z) &= \int_0^z \gamma \exp(-\gamma z) dz \\ &= -\exp(-\gamma z) \Big|_0^z \\ &= -\exp(-\gamma z) - (-\exp(0)) \\ &= 1 - \exp(-\gamma z) \end{aligned}$$

- b. Using your answer to the previous question, evaluate the CDF for the interval from 7 to 12.

Solution:

$$\begin{aligned} F(12) - F(7) &= (1 - \exp(-12\gamma)) - (1 - \exp(-7\gamma)) \\ &= \exp(-7\gamma) - \exp(-12\gamma) \end{aligned}$$

- a. Suppose γ is 3. Given this, what is q , the 10th percentile value of Z?

Solution:

$$\begin{aligned} .1 &= \int_0^q f(z) dz \\ &= F(q) - F(0) \\ .1 &= (1 - \exp(-3q)) - (1 - \exp(0)) \\ .1 &= 1 - \exp(-3q) \\ \exp(-3q) &= .9 \\ -3q &= \log(.9) \\ q &= -\frac{\log(.9)}{3} \\ q &\approx 0.035 \end{aligned}$$

- b. We observe a single random draw from Z , what is the probability this observation is less than .5? Again suppose that $\gamma = 3$.

Solution: We would expect a random draw from this distribution to be less than .5 with a probability equal to .78.

$$\begin{aligned}\int_0^{.5} f(z)dz &= F(.5) - F(0) \\ &= (1 - \exp(-3 \times .5)) - (1 - \exp(0)) \\ &= 1 - \exp(-1.5) \\ &= 0.7768698\end{aligned}$$

Working with normal random variables

Let X and Y be normal random variables with means 0 and 1, respectively, and variances 1 and 4, respectively.³

- a. Find $\Pr(X \leq 1.5)$ and $\Pr(X \leq -1)$.

```
## [1] 0.933
```

```
## [1] 0.159
```

- b. Find the PDF of $(Y - 1)/2$.

The random variable $(Y - 1)/2$ is obtained by subtracting from Y its mean (which is 1) and dividing by the standard deviation (which is 2), so the PDF of $(Y - 1)/2$ is the standard normal.

³BT 3.11