STAT 25100 Lecture 12
5.2 Expectation & Variance of Continuous Random Variables
Section 5.5.1 Hazard Rate Functions

Yibi Huang
Department of Statistics
University of Chicago

Definition: Expected Value of a Continuous R.V.

Let X be a continuous random variable with PDF $f(x)$. The **expected value** or the **expectation** or the **mean** of X, denoted by $\mathrm{E}[X]$, or μ_x is defined to be

$$\mu_x = \mathrm{E}[X] = \int_{-\infty}^{\infty} x\, f(x) \mathrm{d}x$$

provided that provided that $\int_{-\infty}^{\infty} |x| f(x) \mathrm{d}x < \infty$. If the integral diverges, the expectation is undefined.

Integration By Parts (Review)

Since

$$\Big(g(x)h(x)\Big)' = g'(x)h(x) + g(x)h'(x)$$

rearranging the terms

$$g(x)h'(x) = \Big(g(x)h(x)\Big)' - g'(x)h(x)$$

and integrating both sides from a to b, we get

$$\begin{aligned}\int_a^b g(x)h'(x)\mathrm{d}x &= \int_a^b \Big(g(x)h(x)\Big)'\mathrm{d}x - \int_a^b g'(x)h(x)\mathrm{d}x \\ &= \Big[g(x)h(x)\Big]_a^b - \int_a^b g'(x)h(x)\mathrm{d}x\end{aligned}$$

Expected Value — Exponential Distribution

Recall the PDF for Exponential(λ) is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$. The expectation is

$$\mathrm{E}(X) = \int_0^\infty \underbrace{x}_{=g(x)} \cdot \underbrace{\lambda e^{-\lambda x}}_{=h'(x)} \, \mathrm{d}x \quad \rightsquigarrow \begin{cases} g(x) = x & \Rightarrow g'(x) = 1 \\ h'(x) = \lambda e^{-\lambda x} & \Rightarrow h(x) = -e^{-\lambda x} \end{cases}$$

Using integration by parts, we get

$$\begin{aligned} \mathrm{E}(X) = \int_0^\infty \underbrace{x}_{=g(x)} \cdot \underbrace{\lambda e^{-\lambda x}}_{=h'(x)} \, \mathrm{d}x &= \Big[g(x)h(x) \Big]_{x=0}^{\infty} - \int_0^\infty g'(x)h(x)\mathrm{d}x \\ &= \Big[x(-e^{-\lambda x}) \Big]_{x=0}^{\infty} - \int_0^\infty 1 \cdot (-e^{-\lambda x})\mathrm{d}x \\ &= 0 - 0 + \int_0^\infty e^{-\lambda x}\mathrm{d}x \\ &= \frac{-1}{\lambda}(0 - 1) = \frac{1}{\lambda}. \end{aligned}$$

Digression — Odd and Even Functions

A function $g(x)$ is called an *odd function* if it satisfies

$$g(-x) = -g(x), \quad \text{for all } x.$$

A function $h(x)$ is called an *even function* if it satisfies

$$h(-x) = h(x), \quad \text{for all } x.$$

For an odd function $g(x)$,

$$\int_{-\infty}^{0} g(x)\mathrm{d}x \overset{\text{let } x=-y}{=} \int_{0}^{\infty} g(-y)\mathrm{d}y = -\int_{0}^{\infty} g(y)\mathrm{d}y,$$

and hence

$$\int_{-\infty}^{\infty} g(x)\mathrm{d}x = \int_{-\infty}^{0} g(x)\mathrm{d}x + \int_{0}^{\infty} g(x)\mathrm{d}x = -\int_{0}^{\infty} g(x)\mathrm{d}x + \int_{0}^{\infty} g(x)\mathrm{d}x = 0.$$

provided that $\int_{-\infty}^{\infty} |g(x)|\mathrm{d}x < \infty$.

Expected Value — Normal Distribution

Recall the PDF for $X \sim N(\mu,\, \sigma^2)$ is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty.$$

Its expected value is

$$\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{\mu + \sigma z}{\sqrt{2\pi}} e^{-z^2/2} \mathrm{d}z \quad (\text{let } z = \frac{x-\mu}{\sigma} \Rightarrow \mathrm{d}x = \sigma \mathrm{d}z) \\
&= \mu \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \mathrm{d}z}_{=1 \text{ since it's intergal of normal PDF}} + \sigma \underbrace{\int_{-\infty}^{\infty} \frac{z}{\sqrt{2\pi}} e^{-z^2/2} \mathrm{d}z}_{=0 \text{ as } ze^{-z^2/2} \text{ is an odd function}} \\
&= \mu
\end{aligned}$$

If the PDF is an Even function (Symmetric About 0) ...

If the PDF $f(x)$ of a random variable X is an **even** function,

$$f(-x) = f(x) \quad \text{for all } x,$$

then

- $g(x) = xf(x)$ is an odd function since $g(-x) = -xf(-x) = -xf(x) = -g(x)$
- so $\mathrm{E}(X) = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = 0$ provided $\int_{-\infty}^{\infty} |x|f(x)\mathrm{d}x < \infty$.

e.g., the **standard normal** distribution has an expected value of 0 since the PDF is even

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad -\infty < x < \infty,$$

Observe $\phi(-x) = \phi(x)$.

Cauchy Distribution Has No Expected Value

Recall the PDF for Cauchy Distribution is

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

One might think $\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x = 0$ since $f(x)$ is an even function. In fact, its expected value doesn't exist since

$$\int_{-\infty}^{\infty} |x| f(x) \mathrm{d}x = \int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)} \mathrm{d}x = \infty.$$

Calculating $\mathrm{E}(X)$ Using CDF if $X \geq 0$

For a nonnegative continuous random variable X, its expectation can be computed using the CDF $F(x)$ as follows.

$$\mathrm{E}(X) = \int_0^\infty \mathrm{P}(X > x)\mathrm{d}x = \int_0^\infty 1 - F(x)\mathrm{d}x$$

Proof: Using integration by parts, we get

$$\mathrm{E}(X) = \int_0^\infty \underbrace{x}_{=g(x)} \underbrace{f(x)}_{=h'(x)}\, \mathrm{d}x \quad \rightsquigarrow \begin{cases} g(x) = x & \Rightarrow g'(x) = 1 \\ h'(x) = f(x) & \Rightarrow h(x) = F(x) - 1 \end{cases}$$

$$= [g(x)h(x)]_{x=0}^{\infty} - \int_0^\infty g'(x)h(x)\mathrm{d}x$$

$$= [x(F(x) - 1)]_{x=0}^{\infty} - \int_0^\infty 1 \cdot (F(x) - 1)\mathrm{d}x$$

$$= 0 - 0 + \int_0^\infty 1 - F(x)\mathrm{d}x \quad \text{note } \lim_{x\to\infty} x(F(x) - 1) = 0$$

Example

Recall the CDF for $X \sim \text{Exp}(\lambda)$ is

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0 \quad \Rightarrow 1 - F(x) = e^{-\lambda x}.$$

We can use the CDF to find $\mathrm{E}(X)$ the Exponential distribution.

$$\mathrm{E}(X) = \int_0^\infty 1 - F(x)\mathrm{d}x = \int_0^\infty e^{-\lambda x}\mathrm{d}x = \frac{1}{\lambda}.$$

Expected Values for Functions of Random Variables

If X is a continuous random variable with PDF $f_X(x)$, how to find the expected value of $Y = g(X)$?

Method 1: Find the PDF $f_Y(y)$ for Y and calculate its expected value as

$$\mathrm{E}(Y) = \int_{-\infty}^{\infty} y f_Y(y) \mathrm{d}y.$$

Method 2: One can calculate $\mathrm{E}(Y)$ directly using the PDF of X as

$$\mathrm{E}(Y) = \mathrm{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x$$

- Method 2 is easier as it saves the work of finding the PDF for $Y = g(X)$, which is sometimes not easy

Example

Suppose $X \sim \text{Exp}(\lambda)$ with the PDF $f_X(x) = \lambda e^{-\lambda x}$ for $x \geq 0$.
Find the expected value for $Y = e^X$.

Method 1: In Slides `L10_11`, we found the PDF for $Y = e^X$ to be

$$f_Y(y) = \lambda y^{-\lambda-1} \quad \text{for } y \geq 1.$$

$$\mathrm{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y)\mathrm{d}y = \int_{1}^{\infty} y \cdot \lambda y^{-\lambda-1}\mathrm{d}y = \left[\frac{\lambda}{-\lambda+1} y^{-\lambda+1}\right]_{y=1}^{\infty} = \frac{\lambda}{\lambda-1}, \text{ if } \lambda > 1.$$

Method 2: Calculating $\mathrm{E}(Y) = \mathrm{E}(e^X)$ directly using the PDF of X

$$\mathrm{E}(e^X) = \int_{0}^{\infty} e^x f_X(x)\mathrm{d}x = \int_{0}^{\infty} e^x \lambda e^{-\lambda x}\mathrm{d}x = \left[\frac{\lambda}{-(\lambda-1)} e^{-(\lambda-1)x}\right]_{x=0}^{\infty} = \frac{\lambda}{\lambda-1}, \text{ if } \lambda > 1$$

The two methods give identical expected values, but Method 2 takes less work.

Proof for the Equivalence of the Two Methods

We will only prove the case that $g()$ is differentiable & strictly increasing.
Recall in `L10_11`, we showed the PDF of $Y = g(X)$ in this case is

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{d}{dy} g^{-1}(y).$$

So

$$\begin{aligned} \mathrm{E}(Y) = \int_{-\infty}^{\infty} y f_Y(y) \mathrm{d}y &= \int_{-\infty}^{\infty} \underbrace{y}_{=g(x)} \cdot f_X(\underbrace{g^{-1}(y)}_{=x}) \cdot \underbrace{\frac{d}{dy} g^{-1}(y) \mathrm{d}y}_{=\mathrm{d}x} \\ &= \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \mathrm{d}x \end{aligned}$$

where the last equality comes from a change of variables $y = g(x)$, which implies

$$x = g^{-1}(y), \quad \text{and} \quad \mathrm{d}x = \frac{d}{dy} g^{-1}(y) \mathrm{d}y.$$

Variance, Standard Deviation (SD), and Moments

The *variance*, *standard deviation*, *moments*, and *central moments* for a continuous random variable can be defined similarly as for discrete cases.

- Variance (provided that the integral is $< \infty$.):

$$\mathrm{Var}(X) = \sigma^2 = \mathrm{E}\left[(X - \mu_X)^2\right] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) \mathrm{d}x$$

- Shortcut formula for variance: $\mathrm{Var}(X) = \mathrm{E}[(X - \mu)^2] = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$
- Standard deviation (SD): $\mathrm{SD}(X) = \sigma = \sqrt{\mathrm{Var}(X)}$
- kth moment: $\mathrm{E}[X^k]$
- kth central moment: $\mathrm{E}[(X - \mu)^k]$, where $\mu = \mathrm{E}(X)$

provided that $\mathrm{E}[|X|^k] < \infty$ and $\mathrm{E}[|X - \mu|^k] < \infty$.

Remark: Factorial moments for a continuous r.v. can also be defined similarly but they are less useful.

Expected Value and Variance for $aX + b$

If X is a random variable (discrete or continuous), the expected value and the variance for its Linear transformation $Y = g(X) = aX + b$ is

- $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$
- $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$

The proof for the continuous case is omitted since it's similar to the one for the discrete case.

Variance — Normal Distribution

The variance for $X \sim N(\mu, \sigma^2)$ is

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} \mathrm{d}x = \sigma^2 \int_{-\infty}^{\infty} \frac{z^2}{\sqrt{2\pi}} e^{-z^2/2} \mathrm{d}z \quad \left(\begin{array}{l} \text{let } z = \frac{x-\mu}{\sigma} \\ \Rightarrow \mathrm{d}x = \sigma \mathrm{d}z \end{array} \right)$$

It remains to find $\int_{-\infty}^{\infty} z^2 e^{-z^2/2} \mathrm{d}z$. Using integration by parts, we get

$$\int_{-\infty}^{\infty} z^2 e^{-z^2/2} \mathrm{d}z = \int_0^{\infty} \underbrace{z}_{=g(z)} \cdot \underbrace{z e^{-z^2/2}}_{=h'(z)} \mathrm{d}z \rightsquigarrow \begin{cases} g(z) = z & \Rightarrow g'(x) = 1 \\ h'(z) = z e^{-z^2/2} & \Rightarrow h(x) = -e^{-z^2/2} \end{cases}$$

$$= \left[g(z)h(z)\right]_{-\infty}^{\infty} - \int_0^{\infty} g'(z)h(z)\mathrm{d}z$$

$$= \left[z(-e^{-z^2/2})\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} 1 \cdot (-e^{-z^2/2})\mathrm{d}z = 0 - 0 + \int_{-\infty}^{\infty} e^{-z^2/2} \mathrm{d}z = \sqrt{2\pi}.$$

Plugging the above back to $\mathrm{Var}(X)$, we get $\mathrm{Var}(X) = \frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi} = \sigma^2$.

Moments of the Gamma Distribution

Recall PDF for Gamma(α, λ) is $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$, for $x \geq 0$. Its kth moment is

$$\begin{aligned}
\mathrm{E}(X^k) &= \int_0^\infty x^k \cdot \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathrm{d}x = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+k-1} e^{-\lambda x} \mathrm{d}x \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{y}{\lambda}\right)^{\alpha+k-1} e^{-y} \frac{1}{\lambda} \mathrm{d}y \quad (\text{let } y = \lambda x \Rightarrow \mathrm{d}x = \frac{1}{\lambda}\mathrm{d}y) \\
&= \frac{1}{\lambda^k \Gamma(\alpha)} \underbrace{\int_0^\infty y^{\alpha+k-1} e^{-y} \mathrm{d}y}_{=\Gamma(\alpha+k)} = \frac{\Gamma(\alpha+k)}{\lambda^k \Gamma(\alpha)}.
\end{aligned}$$

Using the property $\Gamma(t+1) = t\Gamma(t)$ of the Gamma function, we get

$$\mathrm{E}(X^k) = \frac{\Gamma(\alpha+k)}{\lambda^k \Gamma(\alpha)} = \begin{cases} \alpha/\lambda & \text{if } k = 1 \\ \alpha(\alpha+1)/\lambda^2 & \text{if } k = 2 \\ \alpha(\alpha+1)(\alpha+2)/\lambda^k & \text{if } k = 3 \\ \prod_{i=1}^k (\alpha+k-1)/\lambda^k & \text{in general.} \end{cases}$$

Variance for Gamma Distribution Using Shortcut Formula

In the previous page, we've obtained

$$\mathrm{E}(X) = \frac{\alpha}{\lambda}, \quad \mathrm{E}(X^2) = \alpha(\alpha+1)/\lambda^2.$$

By the shortcut formula, the variance for the Gamma distribution is

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}.$$

It takes more work to calculate $\mathrm{E}(X-\mu)^2 = \mathrm{E}(X-\alpha/\lambda)^2$.

Section 5.5.1 Hazard Rate Functions

Let $X \geq 0$ be the lifetime of some item, with the PDF $f(x)$ and CDF $F(x)$.

Given that the item has survived for a time of t, what's the conditional rate that it will die in the next interval $(t, t + \Delta t)$?

$$
\begin{aligned}
\frac{\mathrm{P}(t < X < t + \Delta t \mid X > t)}{\Delta t} &= \frac{\mathrm{P}(t < X < t + \Delta t, X > t)}{\Delta t \mathrm{P}(X > t)} \\
&= \frac{\mathrm{P}(t < X < t + \Delta t)}{\Delta t \mathrm{P}(X > t)} \\
&= \frac{\int_t^{t+\Delta t} f(x) \mathrm{d}x}{\Delta t (1 - F(t))} \approx \frac{f(t)\Delta t}{\Delta t (1 - F(t))} = \frac{f(t)}{1 - F(t)}
\end{aligned}
$$

The function

$$
\lambda(t) = \frac{f(t)}{1 - F(t)}
$$

is called the *hazard rate function* or *hazard rate*, or *hazard function*.

Example — Exponential Distribution Has a Constant Hazard Rate

Recall the PDF and CDF for Exponential(λ) are

$$f(x) = \lambda e^{-\lambda x}, \quad F(x) = 1 - e^{-\lambda x} \quad \text{for } x \geq 0.$$

The hazard function is thus a constant λ:

$$\lambda(x) = \frac{f(x)}{1 - F(x)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda.$$

Ex: The lifetime of a light bulb has an Exponential with rate $\lambda = 0.001$ per hour.

- The expected lifetime is $1/\lambda = 1/0.001 = 1000$ hours
- Suppose the lightbulb has been on for t hours, the chance that the light bulb dies
  - in the next hour ($\Delta t = 1$ hour) is $\approx \lambda \Delta t = \lambda \cdot 1 = 0.001$
  - in the next minute ($\Delta t = 1/60$ hours) is $\approx \lambda \Delta t = \lambda \cdot (1/60) = 0.001/60$.

Example — Hazard Rate of Weibull Distribution

A continuous random variable X is said to have the **Weibull distribution** with parameters $\beta > 0$ and $\theta > 0$ if it has a PDF of the form

$$f(x) = \frac{\beta}{\theta}(x/\theta)^{\beta-1}\exp(-(x/\theta)^{\beta}), \quad x > 0.$$

Its CDF is

$$F(x) = 1 - \exp(-(x/\theta)^{\beta}), \quad x > 0.$$

The hazard function is thus

$$\lambda(x) = \frac{f(x)}{1-F(x)} = \frac{\frac{\beta}{\theta}(x/\theta)^{\beta-1}\exp(-(x/\theta)^{\beta})}{\exp(-(x/\theta)^{\beta})} = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1}.$$

- The hazard rate increases with x if $\beta > 1$
- The hazard rate decreases with x if $\beta < 1$

Hazard Function Uniquely Determines the CDF

Observe that

$$\frac{d}{\mathrm{d}u}\log(1-F(u)) = \frac{\frac{d}{\mathrm{d}u}(1-F(u))}{1-F(u)} = \frac{-f(u)}{1-F(u)} = -\lambda(u).$$

Integrating both sides from 0 to t, we get

$$-\int_0^t \lambda(u)\mathrm{d}u = [\log(1-F(u))]_{u=0}^t = \log(1-F(t)) - \underbrace{\log(1-\overbrace{F(0)}^{=0})}_{=0} = \log(1-F(t)).$$

Taking exponential on both sides, we get

$$1-F(t) = \exp\left(-\int_0^t \lambda(u)\mathrm{d}u\right) \iff F(t) = 1-\exp\left(-\int_0^t \lambda(u)\mathrm{d}u\right).$$

Example — Hazard of Divorce

The hazard rate $\lambda(t)$ of divorce after t years of marriage is such that

$$\lambda(t) = \frac{1}{2t+1}, \quad t > 0.$$

What's the CDF for $T =$ length of marriage?

Sol:

$$\int_0^t \lambda(u)\mathrm{d}u = \int_0^t \frac{1}{2u+1}\mathrm{d}u = \left[\frac{1}{2}\log(2u+1)\right]_0^t = \frac{1}{2}\log(2t+1) = \log(\sqrt{2t+1}).$$

The CDF is thus

$$F(t) = 1 - \exp\left(-\int_0^t \lambda(u)\mathrm{d}u\right) = 1 - \exp(-\log(\sqrt{2t+1})) = 1 - \frac{1}{\sqrt{2t+1}}, \quad t > 0.$$

The chance that a couple celebrates their 10th wedding anniversary is

$$\mathrm{P}(T > 10) = 1 - F(10) = \frac{1}{\sqrt{2 \times 10 + 1}} = \frac{1}{\sqrt{21}} \approx 0.22.$$