Computer Simulation

Module 2: Calculus, Probability, and Statistics Primers

Dave Goldsman, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

Great Expectations

Lesson Overview

Last Time: We took a brief detour to show how to simulate some easy random variables.

This Time: Back to the formal Probability Review. This module is all about expected values.

Pay particular attention to the LOTUS blossom that you'll soon see!

Definition: The *expected value* (or *mean*) of a RV X is

$$\mathrm{E}[X] \equiv \begin{cases} \sum_x x f(x) & \text{if } X \text{ is discrete} \\ \int_{\mathbb{R}} x f(x)\, dx & \text{if } X \text{ is continuous} \end{cases} = \int_{\mathbb{R}} x\, dF(x).$$

Example: Suppose that $X \sim \text{Bernoulli}(p)$. Then

$$X = \begin{cases} 1 & \text{with prob. } p \\ 0 & \text{with prob. } 1 - p \ (= q) \end{cases}$$

and we have $\mathrm{E}[X] = \sum_x x f(x) = p$. $\square$

Example: Suppose that $X \sim \text{Uniform}(a, b)$. Then

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

and we have $\text{E}[X] = \int_{\mathbb{R}} x f(x)\, dx = (a+b)/2$. $\square$

Example: Suppose that $X \sim \text{Exponential}(\lambda)$. Then

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

and we have (after integration by parts and L'Hôspital's Rule)

$$\mathrm{E}[X] = \int_{\mathbb{R}} x f(x)\, dx = \int_0^\infty x \lambda e^{-\lambda x}\, dx = \frac{1}{\lambda}. \quad \square$$

Def/Thm (the "Law of the Unconscious Statistician" or "LOTUS"): Suppose that $h(X)$ is some function of the RV X. Then

$$\mathrm{E}[h(X)] = \begin{cases} \sum_x h(x) f(x) & \text{if } X \text{ is disc} \\ \int_{\mathbb{R}} h(x) f(x)\, dx & \text{if } X \text{ is cts} \end{cases} = \int_{\mathbb{R}} h(x)\, dF(x).$$

The function $h(X)$ can be anything "nice", e.g., $h(X) = X^2$ or $1/X$ or $\sin(X)$ or $\ell\mathrm{n}(X)$.

Example: Suppose X is the following discrete RV:

x	2	3	4
$f(x)$	0.3	0.6	0.1

Then $\mathrm{E}[X^3] = \sum_x x^3 f(x) = 8(0.3) + 27(0.6) + 64(0.1) = 25.$ $\square$

Example: Suppose $X \sim \mathrm{Unif}(0,2)$. Then

$$\mathrm{E}[X^n] = \int_{\mathbb{R}} x^n f(x)\,dx = 2^n/(n+1). \quad \square$$

Definitions: $\mathrm{E}[X^n]$ is the nth *moment* of X.

$\mathrm{E}[(X - \mathrm{E}[X])^n]$ is the nth *central moment* of X.

$\mathrm{Var}(X) \equiv \mathrm{E}[(X - \mathrm{E}[X])^2]$ is the *variance* of X.

The *standard deviation* of X is $\sqrt{\mathrm{Var}(X)}$.

Theorem: $\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$ (sometimes easier to calculate this way).

Example: Suppose $X \sim \text{Bern}(p)$. Recall that $\text{E}[X] = p$. Then

$$\text{E}[X^2] = \sum_x x^2 f(x) \; = \; p \quad \text{and}$$

$$\text{Var}(X) \; = \; \text{E}[X^2] - (\text{E}[X])^2 \; = \; p(1-p). \quad \Box$$

Example: Suppose $X \sim \text{Exp}(\lambda)$. By LOTUS,

$$\text{E}[X^n] \; = \; \int_0^\infty x^n \lambda e^{-\lambda x}\, dx \; = \; n!/\lambda^n.$$

$$\text{Var}(X) \; = \; \text{E}[X^2] - (\text{E}[X])^2 \; = \; \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 \; = \; 1/\lambda^2. \quad \Box$$

Theorem: $\mathrm{E}[aX+b] = a\mathrm{E}[X]+b$ and $\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$.

Example: If $X \sim \mathrm{Exp}(3)$, then

$$\mathrm{E}[-2X+7] = -2\mathrm{E}[X]+7 = -\frac{2}{3}+7.$$

$$\mathrm{Var}(-2X+7) = (-2)^2\mathrm{Var}(X) = \frac{4}{9}. \quad \square$$

Just a Moment!

We've gotten through most of the good stuff with respect to expectations.

I'll do one more topic – moment generating functions, which are useful for a variety of reasons.

This stuff is a little more challenging, so that's why I've taken a little break.

Definition: $M_X(t) \equiv \mathrm{E}[e^{tX}]$ is the *moment generating function* (mgf) of the RV X. ($M_X(t)$ is a function of t, *not* of X!)

Example: $X \sim \mathrm{Bern}(p)$. Then

$$M_X(t) = \mathrm{E}[e^{tX}] = \sum_x e^{tx} f(x) = e^{t \cdot 1} p + e^{t \cdot 0} q = pe^t + q.$$

Example: $X \sim \mathrm{Exp}(\lambda)$. Then

$$M_X(t) = \int_{\mathbb{R}} e^{tx} f(x)\, dx = \lambda \int_0^\infty e^{(t-\lambda)x}\, dx = \frac{\lambda}{\lambda - t} \text{ if } \lambda > t.$$

Theorem: Under certain technical conditions,

$$\mathrm{E}[X^k] = \frac{d^k}{dt^k} M_X(t) \bigg|_{t=0}, \quad k = 1, 2, \ldots .$$

Thus, you can *generate* the moments of X from the mgf.

Moment generating functions have many other important uses, some of which we'll talk about in this course.

Example: $X \sim \mathrm{Exp}(\lambda)$. Then $M_X(t) = \frac{\lambda}{\lambda - t}$ for $\lambda > t$. So

$$\mathrm{E}[X] = \frac{d}{dt} M_X(t) \bigg|_{t=0} = \frac{\lambda}{(\lambda - t)^2} \bigg|_{t=0} = 1/\lambda.$$

Further,

$$\mathrm{E}[X^2] = \frac{d^2}{dt^2} M_X(t) \bigg|_{t=0} = \frac{2\lambda}{(\lambda - t)^3} \bigg|_{t=0} = 2/\lambda^2.$$

Thus,

$$\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = 1/\lambda^2. \quad \square$$

Summary

Those expectations were certainly great, eh? Especially LOTUS!

Next Time: Suppose you know everything about a random variable. But what can you say if you square it? Or take the log? We'll look at arbitrary functions of RVs. These play important roles in simulation.

Computer Simulation

Module 2: Calculus, Probability, and Statistics Primers

Dave Goldsman, Ph.D.

Professor

Stewart School of Industrial and Systems Engineering

Functions of a Random Variable

GTx

Lesson Overview

Last Time: The lesson certainly had its moments! And mgf's!

This Time: If you know everything about a RV, what can you say about *functions* of the RV?

This has huge implications throughout the course, e.g., in RV generation.

Problem: Suppose we have a RV X with pmf/pdf $f(x)$. Let $Y = h(X)$. Find $g(y)$, the pmf/pdf of Y.

Examples (take my word for it for now):

If $X \sim \text{Nor}(0,1)$, then $Y = X^2 \sim \chi^2(1)$.

If $U \sim \text{Unif}(0,1)$, then $Y = -\frac{1}{\lambda}\ell\text{n}(U) \sim \text{Exp}(\lambda)$.

Discrete Example: Let X denote the number of H's from two coin tosses. We want the pmf for $Y = X^3 - X$.

x	0	1	2
$f(x)$	1/4	1/2	1/4
$y = x^3 - x$	0	0	6

This implies that $g(0) = P(Y = 0) = P(X = 0 \text{ or } 1) = 3/4$ and $g(6) = P(Y = 6) = 1/4$. In other words,

$$g(y) = \begin{cases} 3/4 & \text{if } y = 0 \\ 1/4 & \text{if } y = 6 \end{cases}. \quad \square$$

GTx

Continuous Example: Suppose X has pdf $f(x) = |x|$, $-1 \le x \le 1$. Find the pdf of $Y = X^2$.

First of all, the cdf of Y is

$$\begin{aligned} G(y) &= P(Y \le y) \\ &= P(X^2 \le y) \\ &= P(-\sqrt{y} \le X \le \sqrt{y}) \\ &= \int_{-\sqrt{y}}^{\sqrt{y}} |x|\, dx \;=\; y, \quad 0 < y < 1. \end{aligned}$$

The pdf of Y is $g(y) = G'(y) = 1$, so that $Y \sim \text{Unif}(0, 1)$. $\square$

Inverse Transform Theorem: Suppose X is a continuous random variable having cdf $F(x)$. Then, amazingly, $F(X) \sim \mathrm{Unif}(0,1)$.

Proof: Let $Y = F(X)$. Then the cdf of Y is

$$\begin{aligned} P(Y \le y) &= P(F(X) \le y) \\ &= P(X \le F^{-1}(y)) \\ &= F(F^{-1}(y)) = y, \end{aligned}$$

which is the cdf of the Unif(0,1). $\square$

This result is of great importance when it comes to generating RV's during a simulation.

GTx

Example (how to generate exponential RV's): Suppose $X \sim \text{Exp}(\lambda)$, with cdf $F(x) = 1 - e^{-\lambda x}$ for $x \geq 0$.

So the Inverse Transform Theorem implies that

$$F(X) \;=\; 1 - e^{-\lambda X} \;\sim\; \text{Unif}(0,1).$$

Let $U \sim \text{Unif}(0,1)$ and set $F(X) = U$. Then we have

$$X \;=\; \frac{-1}{\lambda}\,\ell\text{n}(1 - U) \;\sim\; \text{Exp}(\lambda).$$

For instance, if $\lambda = 2$ and $U = 0.27$, then $X = 0.157$ is an Exp(2) realization. $\square$

Exercise: Suppose that X has the Weibull distribution with cdf

$$F(x) = 1 - e^{-(\lambda x)^{\beta}}, x > 0.$$

If you set $F(X) = U$ and solve for X, show that you get

$$X = \frac{1}{\lambda}[-\ell n(1 - U)]^{1/\beta}.$$

Now pick your favorite λ and β, and use this result to generate values of X. In fact, make a histogram of your X values. Are there any interesting values of λ and β you could've chosen?

Bonus Time!

This has been a tough lesson so far.

But if you're stout of heart, you can take a look at one more bonus result…

…or just skip ahead to the Summary. No harm, no foul!

Bonus Theorem: Here's another way to get the pdf of $Y = h(X)$ for some nice continuous function $h(\cdot)$. The cdf of Y is

$$F_Y(y) = P(Y \le y) = P(h(X) \le y) = P(X \le h^{-1}(y)).$$

By the chain rule (and since a pdf must be ≥ 0), the pdf of Y is

$$f_Y(y) = \frac{d}{dy}F_Y(y) = f_X(h^{-1}(y))\left|\frac{d}{dy}h^{-1}(y)\right|.$$

And now, here's how to prove LOTUS!

$$\begin{aligned} \mathrm{E}[Y] &= \int_{\mathbb{R}} y f_Y(y)\,dy = \int_{\mathbb{R}} y f_X(h^{-1}(y)) \left| \frac{d}{dy} h^{-1}(y) \right| dy \\ \text{``=''} &\int_{\mathbb{R}} y f_X(h^{-1}(y))\, dh^{-1}(y) = \int_{\mathbb{R}} h(x) f_X(x)\,dx. \quad \Box \end{aligned}$$

Summary

We looked at the distributions of functions of RVs. Along the way, we used the Inverse Transform Theorem to turn Unif(0,1)'s into arbitrary continuous distributions – very useful in simulation!

Next Time: If you liked one RV, you'll simply *love* two! Joint RVs around the corner…

Computer Simulation

Module 2: Calculus, Probability, and Statistics Primers

Dave Goldsman, Ph.D.

Professor

Stewart School of Industrial and Systems Engineering

Jointly Distributed Random Variables

Lesson Overview

Last Time: Looked at functions of RVs and the Inverse Transform Method for generating RVs from Unif(0,1)'s.

This Time: We'll review two-dimensional RVs that may be correlated with each other. Think height and weight!

In simulation, think consecutive correlated waiting times.

Idea: Consider two random variables interacting together — think height and weight.

Definition: The *joint cdf* of X and Y is

$$F(x,y) \equiv P(X \le x, Y \le y), \quad \text{for all } x, y.$$

Remark: The *marginal cdf* of X is $F_X(x) = F(x, \infty)$. (We use the X subscript to remind us that it's just the cdf of X all by itself.) Similarly, the *marginal cdf* of Y is $F_Y(y) = F(\infty, y)$.

Definition: If X and Y are discrete, then the *joint pmf* of X and Y is $f(x,y) \equiv P(X = x, Y = y)$. Note that $\sum_x \sum_y f(x,y) = 1$.

Remark: The *marginal pmf* of X is

$$f_X(x) = P(X = x) = \sum_y f(x,y).$$

The *marginal pmf* of Y is

$$f_Y(y) = P(Y = y) = \sum_x f(x,y).$$

Example: The following table gives the joint pmf $f(x, y)$, along with the accompanying marginals.

$f(x,y)$	$X = 2$	$X = 3$	$X = 4$	$f_Y(y)$
$Y = 4$	0.3	0.2	0.1	0.6
$Y = 6$	0.1	0.2	0.1	0.4
$f_X(x)$	0.4	0.4	0.2	1

Definition: If X and Y are continuous, then the *joint pdf* of X and Y is $f(x,y) \equiv \frac{\partial^2}{\partial x \partial y} F(x,y)$. Note that $\int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)\, dx\, dy = 1$.

Remark: The *marginal pdf's* of X and Y are

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\, dy \quad \text{and} \quad f_Y(y) = \int_{\mathbb{R}} f(x,y)\, dx.$$

Example: Suppose the joint pdf is

$$f(x,y) = \frac{21}{4}x^2y, \quad x^2 \le y \le 1.$$

Then the marginal pdf's are:

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy = \int_{x^2}^{1} \frac{21}{4}x^2y\,dy = \frac{21}{8}x^2(1-x^4), \ -1 \le x \le 1$$

and

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4}x^2y\,dx = \frac{7}{2}y^{5/2}, \quad 0 \le y \le 1.$$

Definition: X and Y are *independent* RV's if

$$f(x,y) = f_X(x)f_Y(y) \quad \text{for all } x,y.$$

Theorem: X and Y are indep if you can write their joint pdf as $f(x,y) = a(x)b(y)$ for some functions $a(x)$ and $b(y)$, and x and y don't have funny limits (their domains do not depend on each other).

Examples: If $f(x,y) = cxy$ for $0 \le x \le 2$, $0 \le y \le 3$, then X and Y are independent.

If $f(x,y) = \frac{21}{4}x^2y$ for $x^2 \le y \le 1$, then X and Y are *not* independent.

If $f(x,y) = c/(x+y)$ for $1 \le x \le 2$, $1 \le y \le 3$, then X and Y are *not* independent. $\square$

Definition: The *conditional pdf* (or *pmf*) of Y given $X = x$ is $f(y|x) \equiv f(x,y)/f_X(x)$ (assuming $f_X(x) > 0$).

This is a legit pmf/pdf. For example, in the continuous case, $\int_{\mathbb{R}} f(y|x)\,dy = 1$, for any x.

Example: Suppose $f(x,y) = \frac{21}{4}x^2y$ for $x^2 \le y \le 1$. Then

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\frac{21}{4}x^2y}{\frac{21}{8}x^2(1-x^4)} = \frac{2y}{1-x^4}, \quad x^2 \le y \le 1. \quad \square$$

Theorem: If X and Y are indep, then $f(y|x) = f_Y(y)$ for all x, y.

Proof: By definition of conditional and independence,

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)}. \quad \square$$

Definition: The *conditional expectation* of Y given $X = x$ is

$$\mathrm{E}[Y|X = x] \equiv \begin{cases} \sum_y y f(y|x) & \text{discrete} \\ \int_{\mathbb{R}} y f(y|x)\, dy & \text{continuous} \end{cases}$$

Example: The expected weight of a 7' tall guy ($\mathrm{E}[Y|X = 7]$) is > the expected weight of a totally random guy ($\mathrm{E}[Y]$).

Old Cts Example: $f(x, y) = \frac{21}{4} x^2 y$, if $x^2 \le y \le 1$. Then

$$\mathrm{E}[Y|x] = \int_{\mathbb{R}} y f(y|x)\, dy = \int_{x^2}^{1} \frac{2y^2}{1 - x^4}\, dy = \frac{2}{3} \cdot \frac{1 - x^6}{1 - x^4}. \quad \square$$

Summary

Did a lot of stuff: Joint pmf's and pdf's, independent, conditional distributions, and then conditional expectation. Whew!

Next Time: More fun with conditional expectation, including "double trouble"! (Probably the toughest lesson of Bootcamp.)

Computer Simulation

Module 2: Calculus, Probability, and Statistics Primers

Dave Goldsman, Ph.D.

Professor

Stewart School of Industrial and Systems Engineering

Conditional Expectation

Lesson Overview

Last Time: Joint pmf’s and pdf’s, independence, conditional distributions, and a touch of cond’l expectation.

This Time: Continue conditional expectation + several cool applications.

Tough lesson… let’s try to get through it together!

Definition: The *conditional expectation* of Y given $X = x$ is

$$\mathrm{E}[Y|X = x] \equiv \begin{cases} \sum_y y f(y|x) & \text{discrete} \\ \int_{\mathbb{R}} y f(y|x)\, dy & \text{continuous} \end{cases}$$

Example: Suppose $f(x, y) = \frac{21}{4} x^2 y$ for $x^2 \le y \le 1$. Then

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\frac{21}{4} x^2 y}{\frac{21}{8} x^2 (1 - x^4)} = \frac{2y}{1 - x^4}, \quad x^2 \le y \le 1.$$

$$\mathrm{E}[Y|x] = \int_{\mathbb{R}} y f(y|x)\, dy = \int_{x^2}^{1} \frac{2y^2}{1 - x^4}\, dy = \frac{2}{3} \cdot \frac{1 - x^6}{1 - x^4}. \quad \square$$

Double Trouble!

We'll now look at the most-difficult concept of the Probability Bootcamp – *double expectation*.

Idea: the average expected value of all of the conditional expected values is the overall population average.

It's sometimes a very useful tool when calculating such overall averages. Don't Panic!

Theorem (double expectations): $\mathrm{E}[\mathrm{E}(Y|X)] = \mathrm{E}[Y]$.

Proof (cts case): By the Unconscious Statistician,

$$\begin{aligned}
\mathrm{E}[\mathrm{E}(Y|X)] &= \int_{\mathbb{R}} \mathrm{E}(Y|x) f_X(x)\,dx \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} y f(y|x)\,dy \right) f_X(x)\,dx \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} y f(y|x) f_X(x)\,dx\,dy \\
&= \int_{\mathbb{R}} y \int_{\mathbb{R}} f(x,y)\,dx\,dy \\
&= \int_{\mathbb{R}} y f_Y(y)\,dy = \mathrm{E}[Y]. \quad \square
\end{aligned}$$

Old Example: Suppose $f(x,y) = \frac{21}{4}x^2y$, if $x^2 \le y \le 1$. By previous examples, we know $f_X(x)$, $f_Y(y)$, and $\mathrm{E}[Y|x]$. Let's find $\mathrm{E}[Y]$.

Solution #1 (old, boring way):

$$\mathrm{E}[Y] = \int_{\mathbb{R}} y f_Y(y)\,dy = \int_0^1 \frac{7}{2} y^{7/2}\,dy = \frac{7}{9}.$$

Solution #2 (new, exciting way):

$$\begin{aligned}\mathrm{E}[Y] &= \mathrm{E}[\mathrm{E}(Y|X)] = \int_{\mathbb{R}} \mathrm{E}(Y|x) f_X(x)\,dx \\ &= \int_{-1}^{1} \left(\frac{2}{3} \cdot \frac{1-x^6}{1-x^4}\right)\left(\frac{21}{8}x^2(1-x^4)\right) dx = \frac{7}{9}.\end{aligned}$$

Example: A cutesy way to calculate the mean of the Geometric distribution.

Let $Y \sim \text{Geom}(p)$, e.g., Y could be the number of coin flips before H appears, where $P(\text{H}) = p$. From Baby Probability class, we know that the pmf of Y is $f_Y(y) = P(Y = y) = q^{y-1}p$, for $y = 1, 2, \ldots$.

Then the old-fashioned way to calculate the mean is:

$$\text{E}[Y] = \sum_y y f_Y(y) = \sum_{y=1}^{\infty} y q^{y-1} p = 1/p,$$

where the last step follows because I tell you so. $\square$

…Let's use double expectation to do what's called a "standard one-step conditioning argument". Define $X = 1$ if the first flip is H; and $X = 0$ otherwise.

Based on the result X of the first step, we have

$$
\begin{aligned}
\mathrm{E}[Y] &= \mathrm{E}[\mathrm{E}(Y|X)] = \sum_x \mathrm{E}(Y|x) f_X(x) \\
&= \mathrm{E}(Y|X=0)P(X=0) + \mathrm{E}(Y|X=1)P(X=1) \\
&= (1+\mathrm{E}[Y])(1-p) + 1(p). \quad \text{(why?)}
\end{aligned}
$$

Solving, we get $\mathrm{E}[Y] = 1/p$ again! $\square$

Computing Probabilities by Conditioning

Let A be some event, and define the RV $Y = 1$ if A occurs; and $Y = 0$ otherwise. Then

$$\mathrm{E}[Y] = \sum_y y f_Y(y) = P(Y = 1) = P(A).$$

Similarly, for any RV X, we have

$$\mathrm{E}[Y|X = x] = \sum_y y f_Y(y|x) = P(Y = 1|X = x) = P(A|X = x).$$

Thus,

$$\begin{aligned} P(A) &= \mathrm{E}[Y] = \mathrm{E}[\mathrm{E}(Y|X)] \\ &= \int_{\mathbb{R}} \mathrm{E}[Y|X=x]\,dF_X(x) \\ &= \int_{\mathbb{R}} P(A|X=x)\,dF_X(x). \end{aligned}$$

Example/Theorem: If X and Y are independent cts RV's, then

$$P(Y<X) = \int_{\mathbb{R}} P(Y<x)f_X(x)\,dx.$$

Proof: Follows from above result if we let the event $A = \{Y < X\}$.

$\square$

Example: If $X \sim \text{Exp}(\mu)$ and $Y \sim \text{Exp}(\lambda)$ are indep RV's, then

$$\begin{aligned} P(Y < X) &= \int_{\mathbb{R}} P(Y < x) f_X(x)\, dx \\ &= \int_0^{\infty} (1 - e^{-\lambda x}) \mu e^{-\mu x}\, dx \\ &= \frac{\lambda}{\lambda + \mu}. \quad \square \end{aligned}$$

Theorem (variance decomposition):

$$\mathrm{Var}(Y) = \mathrm{E}[\mathrm{Var}(Y|X)] + \mathrm{Var}[\mathrm{E}(Y|X)]$$

Proof (from Ross): By definition of variance and double expectation,

$$\begin{aligned}\mathrm{E}[\mathrm{Var}(Y|X)] &= \mathrm{E}\left[\mathrm{E}(Y^2|X) - \{\mathrm{E}(Y|X)\}^2\right] \\ &= \mathrm{E}(Y^2) - \mathrm{E}\left[\{\mathrm{E}(Y|X)\}^2\right].\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{Var}\left[\mathrm{E}(Y|X)\right] &= \mathrm{E}\left[\{\mathrm{E}(Y|X)\}^2\right] - \{\mathrm{E}\left[\mathrm{E}(Y|X)\right]\}^2 \\ &= \mathrm{E}\left[\{\mathrm{E}(Y|X)\}^2\right] - \{\mathrm{E}(Y)\}^2 .\end{aligned}$$

Thus, putting the last two results together,

$$\mathrm{E}\left[\mathrm{Var}(Y|X)\right] + \mathrm{Var}\left[\mathrm{E}(Y|X)\right] = \mathrm{E}(Y^2) - \{\mathrm{E}(Y)\}^2 = \mathrm{Var}(Y). \quad \square$$

Summary

Studied conditional expectation along with a bunch of its applications, notably, double expectation and the standard conditioning technique.

Next Time: Covariance and correlation.

Congrats! All done with the hardest lesson!

Computer Simulation

Module 2: Calculus, Probability, and Statistics Primers

Dave Goldsman, Ph.D.

Professor

Stewart School of Industrial and Systems Engineering

Covariance and Correlation

Lesson Overview

Last Time: We kicked butt on conditional expectation and its applications.

This Time: We'll talk about independence, covariance, correlation, and related results.

Correlation shows up all over the place in simulation.

GTx

Summary Placeholder

- Talked about blah blah blah…
- This completes Module 2, which went over blah blah blah
- Coming up: Module 3 will blah blah blah