STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-11-21 13:30

the classic density formulae: $W_1 = X + Y$

$$f_{W_1}(w) = \int_{-\infty}^{\infty} f(x, w - x)\, dx$$
$$= \int_{-\infty}^{\infty} f_X(x) f_Y(w - x)\, dx \qquad \text{(when } X \perp Y\text{)}$$

Proof:…

Example: $X \sim N(0, 1)$ and $Y \sim N(0, 1)$ with $X \perp Y$.

Example: $X \sim \mathrm{Unif}[0, 1]$ and $Y \sim \mathrm{Unif}[0, 1]$ with $X \perp Y$.

Exercise (textbook example): $X \sim \mathrm{Exp}(\lambda)$ and $Y \sim \mathrm{Exp}(\lambda)$ with $X \perp Y$

the classic density formulae: $W_2 = Y / X$

$$f_{W_2}(w) = \int_{-\infty}^{\infty} f(x, wx)|x|\,dx$$
$$= \int_{-\infty}^{\infty} f_X(x) f_Y(wx)|x|\,dx \qquad (\text{when } X \perp Y)$$

Proof…

"**Mandatory**" exercise (done in book as example): $X \sim N(0,1)$ and $Y \sim N(0,1)$ with $X \perp Y$. This is a classic. We say W_2 has a Cauchy distribution with density $\frac{1}{\pi}\frac{1}{1+w^2}$

Example: $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\lambda)$ with $X \perp Y$.

the minumum and maximum of n r.v.s

Practical example: A mine has 20 haul trucks with engines whose time-to-failure (from any cause) is random and with Exp(λ) distribution. How long until the first failure of any truck?

Assumption: engines fail independently.

Notation: $X_1, X_2, X_3, \dots, X_{20}$ are independent (assumed) and have the same distribution Exp(λ).

"i.i.d.": *independent and identically distributed*

Nothing special about 20, so let's consider the case of $X_1, \dots, X_n$ i.i.d. Exp(λ). The joint density will be:

$$f(x_1, \dots, x_n) = \lambda e^{-\lambda x_1} \cdots \lambda e^{-\lambda x_n} = \lambda^n e^{-n\lambda \sum x_i}$$

minimum of n i.i.d. exponential r.v.s

Denote by $X_{(1)}$ the time to the first failure.

Theorem: $X_{(1)} \sim \mathrm{Exp}(n\lambda)$.

Proof: …

Exercise (general case): If $X_1, \ldots, X_n$ are i.i.d. each with density $f(x)$ and cdf $F(x)$, the density of $X_{(1)} = \min_{1 \le i \le n}\{X_1, \ldots, X_n\}$ is:

$$f_{X_{(1)}}(x) = nf(x)[1 - F(x)]^{n-1}$$

See section 3.7 up to the end of "E X A M P L E **B**"

Maximum is similar.

expected value

big money!

You and I agree to gamble on the outcome of one toss of one coin.

If H appears, I give you $100. If T appears, you give me $100.

This is a fair game.

Denote by Y my financial outcome. X is discrete with pmf:

$$P(Y = y) = \begin{cases} 0.5 & : y = 100 \\ 0.5 & : y = -100 \end{cases}$$

It's a fair game, so my "average" outcome should be 0. Otherwise it would not be rational for either of us to play the game!

This average is exactly $(100)(0.5) + (-100)(0.5) = 0$.

expected value - discrete case

Definition: The expected value of X that takes on values $\{x_1, x_2, \dots\}$ with pmf $p(x)$ is:

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i)$$

provided $\sum_{i=1}^{\infty} |x_i| p(x_i) < \infty$. Otherwise $E(X)$ is undefined.

some of the "named" distributions

$X \sim \text{Bernoulli}(p)$. Then $E(X) = p$.

Proof: …

$X \sim \text{Binomial}(n, p)$. Then $E(X) = np$.

Proof: …

$X \sim \text{Geometric}(p)$. Then $E(X) = \frac{1}{p}$.

Proof: … (Use: $\frac{d}{dr} \sum_{x=0}^{\infty} r^x = \sum_{x=1}^{\infty} x r^{x-1}$ when $|r| < 1$.)

Exercise (book example): $X \sim \text{Poisson}(\lambda)$. Then $E(X) = \lambda$.

fun with expecations

Suppose X has pmf $p(x) = \frac{6}{\pi^2 x^2}$ on $x \in \{1, 2, 3, \ldots\}$.

Suppose X has pmf $p(x) = \frac{3}{\pi^2 x^2}$ on $x \in \{\pm 1, \pm 2, \pm 3, \ldots\}$.

Treat X as a constant, i.e. $X = a$. Then $E(X) = E(a) = a$.

For a sample space S and event $A \subset S$, consider the "indicator" random variable I_A. Then $E(I_A) = P(A)$.

expected value - continuous case

If X is continuous with density f, its expected value is:

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

provided $\int_{-\infty}^{\infty} |x| f(x)\, dx < \infty$.

Examples: $U \sim \text{Unif}[a, b]$, $Z \sim N(0, 1)$, X Cauchy,…

Exercise (textbook example): $X \sim \text{Gamma}(\alpha, \lambda)$

$E(g(X))$ and extensions

Motivation: suppose $X \sim N(\mu, \sigma^2)$. What is $E(X)$? The answer is μ. Lots of ways to figure this out.

Using the density is tedious but do-able. Or we could use the fact that $X = \mu + \sigma Z$ with $Z \sim N(0, 1)$.

Theorem: Given X and $E(X)$ exists, consider $g(x) = a + bx$. Then $E(g(X)) = E(a + bX) = a + bE(X)$.

Proof: …

$E(g(X))$

A theorem which is too difficult to prove generally is: given X, any* g, and $Y = g(X)$, then:

$$E(Y) = E(g(X)) = \begin{cases} \sum g(x)p(x) & : X \text{ discrete} \\ \int g(x)f(x)\,dx & : X \text{ continuous} \end{cases}$$

in both cases provided the sum/integral congerges "absolutely" (i.e. with $|g(x)|$.)

Example: Average volume of sphere with radius $R \sim \text{Exp}(1)\ldots$

$E(g(X_1, \ldots, X_n))$

Some typical applications:

$$E(X_1 \cdot X_2)$$

$$E\left(X_1 + X_2\right)$$

$$E(X_1 + \cdots + X_n)$$

$$E\left(\overline{X}\right) = E\left(\frac{X_1 + \cdots + X_n}{n}\right)$$

Theorem (continuous version): $X_1, \ldots, X_n$ have joint density $f(x_1, \ldots, x_n)$ and $Y = g(X_1, \ldots, X_n)$. Then:

$$E(Y) = \int \ldots \int g(x_1, \ldots, x_n) f(x_1, \ldots, x_n)\, dx_1 \ldots\, dx_n$$

examples

Suppose $X_1 \perp X_2$. Consider $E(X_1 \cdot X_2)\ldots$

Exercise: $X_1 \perp X_2$. Consider $E(g(X_1)\,h(X_2))$

Now suppose $X_1, \ldots, X_n$ are i.i.d. with $E(X_i) = \mu$. Consider:

$$E\left(\overline{X}\right) = E\left(\frac{X_1 + \cdots + X_n}{n}\right)\ldots$$

$X \sim \mathrm{NegBin}(r, p)\ldots$

putting a number on variation

Expected value is a measure of "location", but random variables with the same "location" can be quite different.

Consider the coin tossing game with $E(Y) = 0$:

$$P(Y = y) = \begin{cases} 0.5 & : y = 100 \\ 0.5 & : y = -100 \end{cases}$$

One thing leads to another. Family trees are compared and contrasted, and after more than a few *schnapps* things get interesting:

$$P(Y_2 = y) = \begin{cases} 0.5 & : y = 1000 \\ 0.5 & : y = -1000 \end{cases}$$

Still, $E(Y_2) = 0$. But the values of Y_2 are more spread out.

variance

One way to measure spread is to use the *variance* of X, defined as:
$\mathrm{Var}(X) = E\left[(X - E(X))^2\right]$.

This is a use of $E(g(X))$ with $g(x) = (x - E(X))^2$.

Very useful:

$$\begin{aligned}
\mathrm{Var}(X) &= E\left(X^2 - 2XE(X) + E(X)^2\right) \\
&= E\left(X^2\right) - 2E(X)E(X) + E(X)^2 \\
&= E\left(X^2\right) - E(X)^2.
\end{aligned}$$

examples

$X \sim \text{Bernoulli}(p) \ldots$

$Z \sim N(0, 1) \ldots$

$X \sim \text{Poisson}(\lambda) \ldots$(uses a trick!)

Variance of $X = a$ constant.

Basic examples for exercise: Exponential, Gamma, Geometric (trick: differentiate power series twice), Binomial (use Poisson trick).

$\mathrm{Var}(a + bX)$, $\mathrm{Var}(X + Y)$ (independent case)

$\mathrm{Var}(a + bX) = b^2\mathrm{Var}(X)$. Proof…

Example: $X \sim N(\mu, \sigma^2)$

When $X \perp Y$, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. Proof…

Actually independence is stronger than necessary. Only needed $E(XY) = E(X)E(Y)$; to be revisited.

variance of the "sample average"

This is a "grand" example of particular importance.

Suppose again $X_1, \ldots, X_n$ is i.i.d. with $E(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.. We already know $E\left(\overline{X}\right) = \mu$.

What about $\mathrm{Var}\left(\overline{X}\right)$?