Math 502AB - Lecture 5

Dr. Jamshidian

September 6, 2017

1 Lecture - Part 1

1.1 Examples: Computing Expectations

1. Suppose that $x \sim gamma(\alpha, \beta)$. Obtain $E(X)$.

   First, note that you do not need to memorize densities (they will be given). To find this expectation, we know that:

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

   In this example, we have:

$$\begin{aligned} E(X) &= \int_0^{\infty} x \left[\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \right] dx \\ &= \int_0^{\infty} \left[\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha} e^{-x/\beta} \right] dx \end{aligned}$$

   Side Note: When talking about **kernel** of a density, we are talking about the part of the density involving the random variable (typically X). The kernel of a density implies that when integrating it over the entire space of the random variable X, you get a constant term. Since a probability distribution requires this integral to be equal to 1, then to get from a *kernel* to a density, you need to divide the kernel by this constant.

   To illustrate:

$$\int_0^{\infty} x^{\alpha-1} e^{-x/\beta} dx = \beta^\alpha \Gamma(\alpha)$$

   which, not coincidentally, is the denominator of the constant term in the PDF of the gamma distribution.

So, back to our example, we have:

$$\frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_{0}^{\infty} x^{\alpha} e^{-x/\beta} dx$$

which is the *kernel* of the distribution $gamma(\alpha+1, \beta)$. Thus we have:

$$E(X) = \frac{\Gamma(\alpha+1)\beta^{\alpha+1}}{\Gamma(\alpha)\beta^{\alpha}} = \frac{\alpha\Gamma(\alpha)\beta}{\Gamma(\alpha)} = \alpha\beta$$

Note: One thing to always keep in mind, is that if we are using the *kernel*, we have to be integrating over the right region.

2. Suppose that $x \sim Poisson(\lambda)$. Obtain $E(X)$.

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!} \\ &= \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{(x-1)!} \end{aligned}$$

There is a potential issue here, since our summation bounds are not the same for the *kernel* as they were before (we had $x = 0$ as the starting value). Thus, we need a change in variables. Letting $y = x - 1$, we have:

$$\begin{aligned} E(X) &= e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y+1}}{y!} \\ &= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!} \\ &= \lambda e^{-\lambda} e^{\lambda} = \lambda \end{aligned}$$

1.2 Expectation of a Function of a R.V.

Theorem: Let X be a random variable and let $Y = g(X)$, for some function g

1. If X is continuous with pdf $f_X(x)$ and $\int_{-\infty}^{\infty} |g(x)| f_X(x) < \infty$, then

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

2. If X is discrete with pmf $f_X(x)$, and $\sum_x |g(x)| f_X(x) < \infty$, then:

$$E[g(x)] = \sum_x g(x) f_X(x)$$

This theorem isn't obvious, but there are large implications. This isn't obvious because if you know the distribution of X, then the distribution of some $g(x)$ might be completely different. Here is an example:

$$\begin{aligned}
Y = g(X) &= -\log(X) \\
X &\sim unif(0,1) \\
Y &\sim exp(1) \\
E(Y) &= \int_0^\infty y e^{-y} dy \\
E(Y) = E[g(X)] &= \int_0^1 -\log x \cdot 1 dx
\end{aligned}$$

Proof: The continuous case requires measure theory. So, let's look at the discrete case:

Let D_x and D_y denote the set of possible values for the random variables X and Y, respectively. Defining $* : E[(g(X)] = \sum_X g(x) f_X(x)$ and $Y = g(X)$, we want to show that $** : \sum_Y y f_Y(y) = *$

$$\begin{aligned}
\sum_{x \in \mathcal{D}_X} g(x) f_X(x) &= \sum_{y \in \mathcal{D}_Y} \sum_{\{x \in \mathcal{D}_X : g(x) = y\}} g(x) f_X(x) \\
&= \sum_{y \in \mathcal{D}_Y} \sum_{\{x \in \mathcal{D}_X : g(x) = y\}} y f_X(x) \\
&= \sum_{y \in \mathcal{D}_Y} y \sum_{\{x \in \mathcal{D}_X : g(x) = y\}} f_X(x) \\
&= \sum_{y \in \mathcal{D}_Y} y P(g(x) = y) \\
&= \sum_{y \in \mathcal{D}_Y} y f_Y(y)
\end{aligned}$$

The key idea in this proof is that, for each y, we are going to get the sum of *all the* x values which map to y.

1.3 Properties of Expectation

1. $E[ax + b] = aE[x] + b$

2. $E[a \cdot g_1(x) + b \cdot g_2(x) + c] = aE[g_1(x)] + bE[g_2(x)] + c$ provided that the expected value of $g_1, g_2 < \infty$

   Proof: The question we're asked is does the sum of two functions with expected values have an expected value itself.

$$\begin{aligned}
&\int_{-\infty}^{\infty} |ag_1(x) + bg_2(x) + c| f_X(x) dx \\
&\leq \int_{-\infty}^{\infty} (|ag_1(x)| + |bg_2(x)| + |c|) \, f_X(x) dx \\
&\text{by triangle inequality}
\end{aligned}$$

   Since each of these component g_i's is finite, the integral is finite and we are done.

3. If $g_1(x) \geq g_2(x) \quad \forall x$, then $E[g_1(x)] \geq E[g_2(x)]$

4. If $a \leq g(x) \leq b \Rightarrow a \leq E[g(x)] \leq b$

Example:

Find the *minimum* over b of $E(X - b)^2$

$$\begin{aligned}
E[X - b]^2 &= E[X - E(X) + E(X) - b]^2 \\
&= E[(X - E(X))^2 + 2(X - E(X))(E(X) - b) + (E(X) - b)^2] \\
&= E[(X - E(X))]^2 + 2E[(X - E(X))(E(X) - b)] + E[(E(X) - b)]^2
\end{aligned}$$

We note that the first term is independent of b and the second term is 0. So, to minimize this equality, it is sufficient to find the value which minimizes the last term. Thus, $b = E(X)$ is the minimum over b.

2 Lecture - Part 2

2.1 Moments and MGF (Section 2.3)

Definition

For a random variable, X, its n^{th} moment is defined by:

$$\mu'_n = E(X^n)$$

And its n^{th} **central moment** is defined by:

$$\mu_n = E[X - E(X)]^n$$

In particular, the second central moment is called the *variance*.

Theorem:

$$Var(X) = E(X^2) - [E(X)]^2$$

This is easy to show by expanding out the term $E[(X - E(X))]^2$.

Examples:

1. Let $X \sim gamma(\alpha, \beta)$. Obtain $Var(X)$.

$$Var(X) = E(X^2) - (E(X))^2 = E(X^2) - \alpha^2\beta^2$$

So to find this result, we need to find $E(X^2)$:

$$\begin{aligned} E(X^2) &= \int_0^\infty x^2 \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha+1} e^{-x/\beta} \\ &= \frac{\Gamma(\alpha+2)\beta^{\alpha+2}}{\Gamma(\alpha)\beta^\alpha} = \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)}\beta^2 = \alpha(\alpha+1)\beta^2 \end{aligned}$$

By noting that the right hand side on the second line is the *kernel* of $gamma(\alpha + 1, \beta)$.

2. Suppose that $X \sim poisson(\lambda)$. Obtain $Var(X)$.

$$Var(X) = E(X^2) - (\lambda)^2$$

So we need to find $E(X^2)$:

$$\begin{aligned} E[X^2] &= \sum_{x=1}^\infty x^2 \frac{e^{-\lambda}\lambda^x}{x!} \\ &= \sum_{x=1}^\infty x \frac{e^{-\lambda}\lambda^x}{(x-1)!} \\ &= \sum_{x=1}^\infty (x-1+1) \frac{e^{-\lambda}\lambda^x}{(x-1)!} \\ &= \sum_{x=1}^\infty (x-1) \frac{e^{-\lambda}\lambda^x}{(x-1)!} + \sum_{x=1}^\infty \frac{e^{-\lambda}\lambda^x}{(x-1)!} \\ &= \lambda \sum_{y=0}^\infty y \frac{e^{-\lambda}\lambda^y}{y!} + \lambda \sum_{x=0}^\infty \frac{e^{-\lambda}\lambda^y}{y!} \\ &= \lambda^2 + \lambda \Rightarrow Var(X) = \lambda^2 \end{aligned}$$

We can do a different approach, that is, using *factorial moments*. In other words, note that:

$$E[X(X-1)] = E(X^2) - E(X)$$

So, we have:

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} \\ &= \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!} \\ &= \lambda^2 \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^{y}}{y!} \\ &= \lambda^2 \end{aligned}$$

Theorem:

$$Var(aX+b) = a^2 Var(X)$$

2.2 Moment Generating Functions

Definition:

Let X be a random variable. Then the *moment generating function* (MGF) of X is denoted by:

$$M_X(t) = E(e^{tX})$$

Provided that the expectation exists for some t in a neighborhood of 0.

Example:

Obtain the *MGF* for $X \sim gamma(\alpha, \beta)$.

$$\begin{aligned} E[e^{tX}] &= \int_0^{\infty} e^{tx}\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1}e^{-x/\beta}dx \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{\infty} x^{\alpha-1}e^{-x(\frac{1}{\beta}-t)}dx \end{aligned}$$

This is the kernel of $gamma(\alpha, \frac{\beta}{1-t\beta})$. But what is the condition under which this integral will be finite? We know both parameters have to be positive, so:

$$E[e^{tX}] = \frac{1}{\Gamma(\alpha)\beta^\alpha}\left[\Gamma(\alpha)\left(\frac{\beta}{1-t\beta}\right)\right], \quad \text{for } t < \frac{1}{\beta}$$

Theorem:

If $M_X(t)$ is the *moment generating function* of X, then:

$$M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(t) = E[X^n], \text{ when evaluated at } t = 0$$

To show this:

$$\begin{aligned}
\frac{d^n}{dt^n} M_X(t) &= \frac{d^n}{dt^n} E\left[e^{tX}\right] \\
&= E\left[\frac{d^n}{dt^n} e^{tX}\right] \\
&= E\left[X^n e^{tX}\right]
\end{aligned}$$

This implies that the n^{th} derivative of the moment generating function evaluated at 0 is $E(X^n)$.

Example:

1. Consider $X \sim gamma(\alpha, \beta)$. Use the MGF to obtain $E(X)$ and $Var(X)$.

$$M_X(t) = \left(\frac{1}{1 - t\beta}\right)^{\alpha}$$

We then have:

$$\begin{aligned}
M'_X(t) &= -\alpha(1 - t\beta)^{-\alpha-1}(-\beta) \\
M'_X(0) &= \alpha\beta
\end{aligned}$$

$$\begin{aligned}
M''_X(t) &= \alpha\beta(-\alpha - 1)(1 - t\beta)^{-\alpha-2}(-\beta) \\
E(X^2) = M''_X(0) &= \alpha(\alpha + 1)\beta^2 \\
&\Rightarrow Var(X) = \alpha\beta^2
\end{aligned}$$

2. Obtain the MGF for $X \sim poisson(\lambda)$:

$$\begin{aligned}
E[e^{tX}] &= \sum_{x=0}^{\infty} e^{tX} \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\lambda}(\lambda e^t)^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{-\lambda} \cdot e^{\lambda e^t} = e^{\lambda e^t - \lambda}
\end{aligned}$$