STA 623 – Fall 2013 – Dr. Charnigo

Section 3.6: Inequalities and Identities

Chebychev's Inequality. We more or less proved Chebychev's Inequality earlier this semester (Section 2.3) when showing that a random variable X for which $Var[X] = 0$ necessarily had $X = E[X]$ with probability one. As a refresher, for any positive integer j we have

$$\begin{aligned} E[(X - E[X])^2] &\geq E[(X - E[X])^2 1_{\{|X - E[X]| \geq 1/j\}}] \\ &\geq E[(1/j)^2 1_{\{|X - E[X]| \geq 1/j\}}] \\ &\geq (1/j)^2 P(|X - E[X]| \geq 1/j). \end{aligned}$$

Replacing $1/j$ by a positive number ϵ and rearranging, we have

$$P(|X - E[X]| \geq \epsilon) \leq \epsilon^{-2} E[(X - E[X])^2].$$

In fact, Chebychev's Inequality can be made more general. We can replace $(X - E[X])^2$ by any nonnegative $g(X)$ with finite expectation:

$$P(g(X) \geq \epsilon^2) \leq \epsilon^{-2} E[g(X)].$$

Unfortunately, Chebychev's Inequality is extremely conservative. For instance, if $g(X) = (X - E[X])^2$ and $\epsilon^2 = Var[X]$, then we obtain

$$P(|X - E[X]| \geq SD[X]) \leq 1,$$

a true but manifestly useless statement. Part of the problem is that Chebychev's Inequality does not exploit any information about the distribution of X (other than that it has finite mean and variance). Inequalities that exploit information about the distribution of X are usually less conservative.

A moment generating inequality. Here is an interesting inequality: at any $t \geq 0$ for which $M_X(t)$ is finite, and for any real number a, we have

$$P(X \geq a) \leq \exp[-at] M_X(t).$$

Let's prove this result. Assume for simplicity that X is a continuous random variable with probability density function $f(x)$. Then

$$\begin{aligned} P(X \geq a) &= \int_a^\infty f(x)\,dx \\ &\leq \int_a^\infty \exp[(x-a)t] f(x)\,dx \\ &\leq \int_{-\infty}^\infty \exp[(x-a)t] f(x)\,dx \\ &= \exp[-at] M_X(t). \end{aligned}$$

The first "$\leq$" holds because and the second "$\leq$" holds because

As an application, let us find a bound for the probability with which a gamma random variable X with parameters $\alpha \in (0,\infty)$ and $\beta \in (0,\infty)$ exceeds its mean by more than one standard deviation. Taking for granted that $E[X] = \alpha\beta$, $Var[X] = \alpha\beta^2$, and $M_X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha$ for $|t| < 1/\beta$, by putting $a := E[X] + SD[X] = \alpha\beta + \sqrt{\alpha}\beta$ and $t := 1/(C\beta)$ with $C \in (1,\infty)$ we have

$$\begin{aligned} P(X \geq \alpha\beta + \sqrt{\alpha}\beta) &\leq \exp[-(\alpha\beta + \sqrt{\alpha}\beta)/(C\beta)] \left(\frac{1}{1-1/C}\right)^\alpha \\ &= \exp[-\alpha/C - \sqrt{\alpha}/C + \alpha\log C - \alpha\log(C-1)]. \end{aligned}$$

The question then becomes, given $\alpha \in (0,\infty)$, for which $C \in (1,\infty)$ is the bound smallest? Since the exponential function is monotone increasing in its argument, we can answer this question by maximizing $\alpha/C + \sqrt{\alpha}/C - \alpha\log C + \alpha\log(C-1)$ with respect to $C \in (1,\infty)$. This is a calculus exercise; the end result is $C = \sqrt{\alpha} + 1$. With this choice of C we have

$$P(X \geq \alpha\beta + \sqrt{\alpha}\beta) \leq \exp[-\alpha/(\sqrt{\alpha}+1) - \sqrt{\alpha}/(\sqrt{\alpha}+1) + \alpha\log(\sqrt{\alpha}+1) - \alpha\log(\sqrt{\alpha})].$$

For example, if $\alpha = 1$, then we obtain a bound of $\exp[-1/2 - 1/2 + \log 2] = 0.7358$. This bound is very conservative (the actual probability is 0.1353) but not as bad as the useless bound of 1 provided by Chebychev's Inequality. Note that I am identifying $\{X \geq E[X] + SD[X]\}$ with $\{|X - E[X]| \geq SD[X]\}$. Is that justified?

A gamma identity. Your textbook authors list several identities. None of these identities is particularly memorable, but the technique used to prove them is worth learning.

As an illustration, I will prove an identity that your authors did not, namely that if X is a gamma random variable with parameters $\alpha \in (0, \infty)$ and $\beta \in (0, \infty)$ then

$$E[g(X)(X - \alpha\beta)] = \beta E[Xg'(X)]$$

whenever $g(x)$ is a continuously differentiable function on $(0, \infty)$ for which both of the expectations exist as finite numbers and $\lim_{x\to\infty} g(x)xf(x) = \lim_{x\to 0} g(x)xf(x) = 0$. (Why stipulate that the latter limit be 0? Aren't we covered since $\lim_{x\to 0} x = 0$?)

Putting

$$u := \beta x f(x) = \frac{\beta^{-\alpha+1}}{\Gamma[\alpha]} x^{\alpha} \exp[-x/\beta]$$

and $dv := g'(x)\ dx$ for integration by parts, we obtain $v = g(x)$ and

$$du = \frac{\beta^{-\alpha+1}}{\Gamma[\alpha]} \exp[-x/\beta](\alpha x^{\alpha-1} - x^{\alpha}/\beta)\ dx,$$

so that

$$\begin{aligned}
\beta E[Xg'(X)] &= \int_0^{\infty} \frac{\beta^{-\alpha+1}}{\Gamma[\alpha]} x^{\alpha} \exp[-x/\beta] g'(x)\ dx \\
&= [g(x)\beta x f(x)]_0^{\infty} - \int_0^{\infty} g(x) \frac{\beta^{-\alpha+1}}{\Gamma[\alpha]} \exp[-x/\beta](\alpha x^{\alpha-1} - x^{\alpha}/\beta)\ dx \\
&= -\int_0^{\infty} \\
&= \int_0^{\infty} \\
&= E[g(X)(X - \alpha\beta)].
\end{aligned}$$