3.6 Inequalities and Identities

Theorem 3.6.1 (Chebychev's Inequality)

Let X be a random variable and let $g(x)$ be a nonnegative function. Then, for any $r > 0$,

$$P(g(X) \geq r) \leq \frac{Eg(X)}{r}.$$

Proof:

$$\begin{aligned} Eg(X) &= \int_{-\infty}^{\infty} g(x) f_X(x) dx \\ &\geq \int_{\{x: g(x) \geq r\}} g(x) f_X(x) dx \quad (g \text{ is nonnegative}) \\ &\geq r \int_{\{x: g(x) \geq r\}} f_X(x) dx \\ &= rP(g(X) \geq r). \end{aligned}$$

Rearranging now produces the desired inequality. $\square$

Example 3.6.2 (Illustrating Chebychev)

let $g(x) = (x-\mu)^2/\sigma^2$, where $\mu = \mathrm{E}X$ and $\sigma^2 = \mathrm{Var}X$. For convenience write $r = t^2$. Then

$$P(\frac{(X-\mu)^2}{\sigma^2} \geq t^2) \leq \frac{1}{t^2} E \frac{(X-\mu)^2}{\sigma^2} = \frac{1}{t^2}.$$

Thus,

$$P(|X-\mu| \geq t\sigma) \leq \frac{1}{t^2}.$$

For example, taking $t = 2$, we get

$$P(|X-\mu| \geq 2\sigma) \leq \frac{1}{2^2} = 0.25.$$

Example 3.6.3 (A normal probability inequality)

If Z is standard normal, then

$$P(|Z| \geq t) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}, \quad \text{for all } t > 0.$$

Write

$$\begin{aligned} P(Z \geq t) &= \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2} dx \\ &\leq \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{x}{t} e^{-x^2/2} dx \quad (\text{since } x/t > 1) \\ &= \frac{1}{\sqrt{2\pi}} \frac{e^{-t^2/2}}{t} \end{aligned}$$

and use the fact that $P(|Z| \geq t) = 2P(Z \geq t)$.

Theorem 3.6.4

Let $X_{\alpha,\beta}$ denote a gamma(α, β) random variable with pdf $f(x|\alpha, \beta)$, where $\alpha > 1$. Then for any constants a and b,

$$P(a < X_{\alpha,\beta} < b) = \beta f(a|\alpha, \beta) - f(b|\alpha, \beta) + P(a < X_{\alpha-1,\beta} < b).$$

Lemma 3.6.5(Stein's Lemma)

Let $X \sim N(\theta, \sigma^2)$, and let g be a differentiable function satisfying $E|g'(X)| < \infty$. Then

$$E[g(X)(X - \theta)] = \sigma^2 Eg'(X).$$

Proof: The left-hand side is

$$E[g(X)(X - \theta)] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} g(x)(x - \theta) e^{-(x-\theta)^2/(2\sigma^2)} dx.$$

Using integration by parts with $u = g(x)$ and $dv = (x - \theta)e^{-(x-\theta)^2/(2\sigma^2)} dx$ to get

$$E[g(X)(X - \theta)] = \frac{1}{\sqrt{2\pi}\sigma} \big[- \sigma^2 g(x) e^{-(x-\theta)^2/(2\sigma^2)} |_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} g'(x) e^{-(x-\theta)^2/(2\sigma^2)} dx \big].$$

The condition on g' is enough to ensure that the first term is 0 and what remains on the right-hand side is $\sigma^2 Eg'(X)$. $\square$

Example 3.6.6 (Higher-order normal moments)

Stein's lemma makes calculation of higher-order moments quite easy/ For example, if $X \sim$

$N(\theta, \sigma^2)$, then

$$\begin{aligned} EX^3 &= EX^2(X - \theta + \theta) = EX^2(X - \theta) + \theta EX^2 \\ &= 2\sigma^2 EX + \theta EX^2 = 2\sigma^2\theta + \theta(\sigma^2 + \theta^2) \\ &= 3\theta\sigma^2 + \theta^3. \end{aligned}$$

<u>Theorem 3.6.7</u>

Let χ_p^2 denote a chi-squared random variable with p degrees of freedom. For any function $h(x)$,

$$Eh(\chi_p^2) = pE\big(\frac{h(\chi_{p+2}^2)}{\chi_{p+2}^2}\big)$$

provided the expectations exist.

Some moment calculations are very easy with Theorem **??**. For example, the mean of a χ_p^2 is

$$E\chi_p^2 = pE\big(\frac{\chi_p^2}{\chi_p^2}\big) = pE(1) = p,$$

and the second moment is

$$E(\chi_p^2)^2 = pE\big(\frac{(\chi_p^2)^2}{\chi_p^2}\big) = pE(\chi_p^2) = p(p+2).$$

So $\mathrm{Var}\big(\chi_p^2\big) = p(p+2) - p^2 = 2p$.

<u>Theorem 3.6.8</u> (Hwang)

Let $g(x)$ be a function with $-\infty < Eg(X) < \infty$ and $-\infty < g(-1) < \infty$. Then:

a. If $X \sim Poisson(\lambda)$,

$$E(\lambda g(X)) = E(Xg(X-1)).$$

b. If $X \sim$ negative binomial(r, p),

$$E((1-p)g(X)) = E\big(\frac{X}{r + X - 1} g(X-1)\big).$$

<u>Example 3.6.9</u> (Higher-order Poisson moments)

For $X \sim$ Poisson(λ), take $g(x) = x^2$ and use Theorem 3.6.8:

$$E(\lambda X^2) = E(X(X-1)^2) = E(X^3 - 2X^2 + X).$$

Therefore, the third moment of a Poisson(λ) is

$$\begin{aligned} \mathrm{E}X^3 = \lambda \mathrm{E}X^2 &= 2\mathrm{E}X^2 - \mathrm{E}X \\ &= \lambda(\lambda + \lambda^2) + 2(\lambda + \lambda^2) - \lambda = \lambda^3 + 3\lambda^2 + \lambda. \end{aligned}$$

For the negative binomial, the mean can be calculated by taking $g(x) = r + x$,

$$\mathrm{E}((1-p)(r+X)) = \mathrm{E}\Big(\frac{X}{r+X-1}(r+X-1)\Big) = \mathrm{E}X,$$

so, rearranging, we get

$$\mathrm{E}X = \frac{r(1-p)}{p}.$$