3.5 Inequalities and identities

When we can't calculate probabilities, it is important to estimate them by inequalities to obtain bounds.

Theorem (Chebychev's inequality)

Let X be a R.V.

$$g(x) \geq 0 \quad \forall x$$

Then

$$P(g(X) \geq r) \leq \frac{\mathbb{E} g(X)}{r} \qquad \text{for all } r > 0.$$

Proof.

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

$$\geq \int_{x : g(x) \geq r} g(x) f_X(x)\, dx$$

$$\geq \int_{x : g(x) \geq r} r f_X(x)\, dx$$

$$= r\, P(g(X) \geq r)$$

Dividing both sides by r to conclude. $\square$

Example.

Let $g(x) = \dfrac{(x-\mu)^2}{\sigma^2}$, where $\begin{cases} \mu = \mathbb{E}X \\ \sigma^2 = \operatorname{var} X \end{cases}$

Then $P\left(\dfrac{(x-\mu)^2}{\sigma^2} \geqslant r\right) \leqslant \dfrac{1}{r} \mathbb{E} \dfrac{(x-\mu)^2}{\sigma^2}$

$$= \frac{1}{r}$$

Write $r = t^2$, $t > 0$ to obtain

$$P(|x-\mu| \geqslant t\sigma) \leqslant \frac{1}{t^2}.$$

$$P(|x-\mu| < t\sigma) \geqslant 1 - \frac{1}{t^2}.$$

Take $t = 2$: $P(|x-\mu| \geqslant 2\sigma) \leqslant 1/4 = 25\%$.

$t = 3$: $P(|x-\mu| \geqslant 3\sigma) \leqslant 1/9 = 11.1..\%$

$t = 4$: $P(|x-\mu| \geqslant 4\sigma) \leqslant 1/16 = 6.25\%$

and so on.

Theorem (tighter inequality for normal tails)

if $Z \sim N(0,1)$ then for $t>0$:

$$P(|Z| \geq t) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}$$

Remark.

if $t = 2$, $P(|Z| \geq 2) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-2}}{2} = 0.054 < 1/4$ (above)

in general

$$\underbrace{\sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}}_{\text{exponentially small}} \ll \underbrace{\frac{1}{t^2}}_{\text{polynomially small}}$$

Proof.

$$P(|Z| \geq t) = 2\, P(Z \geq t) \quad \leftarrow f_Z \text{ symmetric}$$

$$= 2 \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$\leq 2 \int_t^\infty \frac{1}{\sqrt{2\pi}} \frac{z}{t} e^{-\frac{1}{2}z^2} dz$$

$$= \frac{2}{\sqrt{2\pi}} \frac{1}{t} \left. e^{-\frac{1}{2}z^2} \right|_\infty^t$$

$$= \sqrt{\frac{2}{\pi}} \frac{1}{t} e^{-\frac{1}{2}t^2}. \qquad \square$$

Identities

$\boxed{\text{Poisson}}$: if $X \sim \text{Poisson}(\lambda)$

then

$$f_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

So

$$\begin{cases} P(X=0) = e^{-\lambda} \\ P(X=x+1) = e^{-\lambda} \dfrac{\lambda^{x+1}}{(x+1)!} \end{cases}$$

$$= P(X=x) \frac{\lambda}{x+1}.$$

Recursion-like identities such as this may be useful in various situations that require such computations

Gamma: if $X_{\alpha,\beta} \sim \text{Gamma}(\alpha,\beta)$ with pdf $f(x|\alpha,\beta)$, $\alpha > 1$, $\beta > 0$

Then $\forall a, b$

$$\boxed{P(X_{\alpha,\beta} \in (a,b)) = \beta\left(f(a|\alpha,\beta) - f(b|\alpha,\beta)\right) + P(X_{\alpha-1,\beta} \in (a,b))}$$

Remarks.

- Proof is an application of integration by parts (exercise!)
- if $\alpha \in \mathbb{N}$, the above identity allows us to recurse to $\text{Gamma}(\alpha-1,\beta)$; $\text{Gamma}(\alpha-2,\beta)$, ..., and so on to $\text{Gamma}(1,\beta) \equiv \text{Exp}(\beta)$.

Stein's identity for Normal variables

if $X \sim N(\theta, \sigma^2)$

g a differentiable function s.t. $\mathbb{E}|g'(X)| < \infty$.

Then

$$\boxed{\mathbb{E}\left[g(X)(X-\theta)\right] = \sigma^2 \, \mathbb{E}\, g'(X).}$$

Proof.

$$\text{LHS} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} g(x)(x-\theta)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2}\, dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} g(x)(-\sigma^2)\, d\left(e^{-\frac{1}{2\sigma^2}(x-\theta)^2}\right)$$

$$\overset{\text{i.b.p.}}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \left(g(x)(-\sigma^2)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \sigma^2 e^{-\frac{1}{2\sigma^2}(x-\theta)^2}\, dg(x) \right)$$

$\to 0$

$(*)$ see next page

$$= \sigma^2\, \mathbb{E}\, g'(X). \qquad \square$$

Verifying $(*)$: Need to check $g(x)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \to 0$ as $x \to \infty$ (and Similarly, as $x \to -\infty$)

Fix $x_0 > \theta$.

$$g(x) = g(x_0) + \int_{x_0}^{x} g'(t)\, dt$$

So
$$g(x)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2} = \underbrace{e^{-\frac{1}{2\sigma^2}(x-\theta)^2} g(x_0)} + \int_{x_0}^{x} g'(t)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2}\, dt$$

$$\Rightarrow \left| g(x)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \right| \leq A(x) + \int_{x_0}^{x} |g'(t)|\, e^{-\frac{1}{2\sigma^2}(t-\theta)^2}\, dt$$

Since $(t-\theta)^2 \leq (x-\theta)^2$ if $t \in (x_0, x)$

$$\leq A(x) + \underbrace{\int_{x_0}^{\infty} |g'(t)|\, e^{-\frac{1}{2\sigma^2}(t-\theta)^2}\, dt}_{B(x_0)} \qquad (1)$$

Note that (1) holds for any $x > x_0 > \theta$

Let $x \to \infty$ then $A(x) \to 0$

Thus
$$\limsup_{x \to \infty} \left| g(x)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \right| \leq B(x_0) \quad \forall x_0 > \theta.$$

Now, let $x_0 \to \infty$ then $B(x_0) \to 0$ because $\mathbb{E}\,|g'(X)| < \infty$ and Lebesgue's dominated convergence theorem.

Hence,
$$\limsup_{x \to \infty} \left| g(x)\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \right| = 0 \quad . \ \square$$

Application.

we know $\mathbb{E}X = \theta$

$$\mathbb{E}X^2 = (\mathbb{E}X)^2 + \operatorname{var} X$$
$$= \theta^2 + \sigma^2$$

How about $\mathbb{E}X^3$?

By Stein's lemma

$$\mathbb{E}X^3 = \mathbb{E}\, X^2(X-\theta+\theta)$$
$$= \mathbb{E}\, X^2(X-\theta) + \theta\, \mathbb{E}X^2$$

$g(x) = x^2$

$$= \sigma^2\, \mathbb{E}(2X) + \theta(\theta^2+\sigma^2)$$

$$= 2\sigma^2\theta + \theta^3 + \theta\sigma^2$$
$$= \theta^3 + 3\sigma^2\theta.$$

and so on, for $\mathbb{E}X^n$, $n = 4, 5, \ldots$

Hwang's identities for some discrete variables

Let g be a function with $|\mathbb{E}\, g(X)| < \infty$.
and $|g(-1)| < \infty$.

- if $X \sim \text{Poisson}(\lambda)$ then

$$\boxed{\mathbb{E}\, \lambda g(X) = \mathbb{E}\, X g(X-1).}$$

- if $X \sim \text{NegBinomial}(r,p)$ then

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

Application.

- if $X \sim \text{Poisson}(\lambda)$ then $\mathbb{E}X = \text{var}X = \lambda$

Thus $\mathbb{E}\, \lambda X^2 = \mathbb{E} X (X-1)^2 \leftarrow g(x) = x^2$

$$\lambda(\lambda^2+\lambda) = \mathbb{E}(X^3 - 2X^2 + X)$$
$$\lambda^3 + \lambda^2 = \mathbb{E}X^3 - 2(\lambda^2+\lambda) + \lambda$$
$$\Rightarrow \quad \mathbb{E}X^3 = \lambda^3 + 3\lambda^2 + \lambda.$$

- if $X \sim \text{NegBinomial}(r,p)$

Take $g(x) = r+x$ then

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

$$\Rightarrow \quad (1-p)r + (1-p)\,\mathbb{E}X = \mathbb{E}X$$

$$\Rightarrow \quad \mathbb{E}X = \frac{(1-p)r}{p}.$$