Some Useful Inequalities and Identities

Theorem 5.1 (Chebychev's Inequality)

Let X be a random variable and $g(x)$ be a non-negative function such that $\mathsf{E}(g(X))$ *exists. Then, for any constant* $r > 0$,

$$\mathsf{P}\Big(g(X) \geqslant r\Big) \;\leqslant\; \frac{\mathsf{E}\left(g(X)\right)}{r}$$

Corollary 5.1.1

Let X be a random variable with finite mean μ and variance σ^2 then for any $t > 0$

$$\mathsf{P}\Big(|X-\mu| \geqslant t\sigma\Big) \;\leqslant\; \frac{1}{t^2}$$

Markov's Inequality

Theorem 5.2 (Markov's Inequality)
Suppose X *is a non-negative random variable with* $\mathsf{P}(X = 0) < 1$ *then for any constant* $r > 0$

$$\mathsf{P}(X \geqslant r) \;\leqslant\; \frac{\mathsf{E}(X)}{r}$$

* In fact, Markov showed that equality in the above theorem is attained if, and only if, X takes probabilities on only 2 points 0 and r which clearly shows that Chebychev's bound is almost never attained.

Normal Tail Probabilities

* If we restrict to the normal family of distributions then we get the following useful result.

* The first part gives much tighter bounds than Chebychev for this distribution.

* The second part gives us a lower bound on the probabilities which cannot be found in general.

Theorem 5.3

If Z is a standard normal random variable then

$$\mathsf{P}\Big(|Z| \geqslant t\Big) \leqslant \sqrt{\frac{2}{\pi}}\frac{\mathrm{e}^{-t^2/2}}{t}$$

$$\mathsf{P}\Big(|Z| \geqslant t\Big) \geqslant \sqrt{\frac{2}{\pi}}\frac{t\mathrm{e}^{-t^2/2}}{1+t^2}$$

Another Inequality Related to Chebychev's

* When moment generating functions exist we get the following result.

* Depending on the value of t used, we can often get better bounds with this also.

Theorem 5.4

Suppose that X is a random variable whose moment generating function $M_X(t)$ exists for $-h < t < h$. Then for any constant $a \in \mathbb{R}$,

$$\begin{aligned} \mathrm{P}(X \geqslant a) &\leqslant \mathrm{e}^{-at} M_X(t) && \text{for any } 0 < t < h \\ \mathrm{P}(X \leqslant a) &\leqslant \mathrm{e}^{-at} M_X(t) && \text{for any } -h < t < 0 \end{aligned}$$

Recurrence Relationships

* For discrete random variables we can often write recurrence relationships of the form

$$\mathsf{P}(X = x + 1) \; = \; h\Big(\mathsf{P}(X = x)\Big)$$

for some function h.

* These relationships are called recurrence relationships.

* If there is a certain x for which the probability is known or very easily calculated, these relationships can make other probabilities easy to find also.

Some Discrete Recurrence Relationships

* If X is a Poisson random variable with mean λ then

$$\mathsf{P}(X = x + 1) \; = \; \frac{\lambda}{x+1}\mathsf{P}(X = x) \qquad x = 0, 1, 2, \ldots$$

* If $X \sim$ Binomial(n, p) then

$$\mathsf{P}(X = x{+}1) \; = \; \frac{p(n-x)}{(1-p)(x+1)}\mathsf{P}(X = x) \qquad x = 0, 1, \ldots, n{-}1$$

* If $X \sim$ Negative Binomial(r, p) then

$$\mathsf{P}(X = x + 1) \; = \; \frac{(1-p)(r+x)}{x+1}\mathsf{P}(X = x) \qquad x = 0, 1, 2, \ldots$$

A Relationship for the Gamma Distribution

* Gamma probabilities are hard to find in general.

* If the shape parameter is a positive integer, however, they can be found by integration by parts.

* In this case the recurrence has to do with probabilities for different random variables in the same family.

* Since probabilities are easy to find for the case of $\alpha = 1$ we can use these relationships to find them for any positive integer α.

A Relationship for the Gamma Distribution

Theorem 5.5

Suppose that $X_{\alpha,\beta}$ *is a gamma random variable with probability density function*

$$f(x \mid \alpha, \beta) \;=\; \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} \mathrm{e}^{-x/\beta} \qquad x > 0$$

Suppose $\alpha > 1$ *and* a *and* b *are any non-negative constants with* $a < b$ *then*

$$\mathrm{P}\left(a < X_{\alpha,\beta} < b\right) \;=\; \beta\Big(f(a \mid \alpha, \beta) - f(b \mid \alpha, \beta)\Big) + \mathrm{P}\left(a < X_{\alpha-1,\beta} < b\right)$$

Corollary 5.5.1

Let $f(x \mid \alpha, \beta)$ *and* $F(x, \mid \alpha, \beta)$ *denote the pdf and cdf of a Gamma*(α, β) *random variable. Then for* $\alpha > 1$ *and any* $x > 0$

$$F(x \mid \alpha, \beta) \;=\; F(x \mid \alpha - 1, \beta) - \beta f(x \mid \alpha, \beta)$$

Stein's Lemma

Theorem 5.6 (Stein's Lemma)

Suppose that Z *is a standard normal random variable and let* $g(x)$ *be a differentiable function such that* $\mathsf{E}\left(|g'(Z)|\right) < \infty$ *then*

$$\mathsf{E}\Big(Zg(Z)\Big) \;=\; \mathsf{E}\Big(g'(Z)\Big)$$

Corollary 5.6.1

Suppose that $X \sim$ *Normal*(μ, σ^2) *and let* $g(x)$ *be a differentiable function such that* $\mathsf{E}\left(|g'(X)|\right) < \infty$ *then*

$$\mathsf{E}\Big((X-\mu)g(X)\Big) \;=\; \sigma^2\,\mathsf{E}\Big(g'(X)\Big)$$

Other Useful Identities for Distributions

Theorem 5.7

Let X_p be a χ^2_p random variable and let h be a function whose expectation exists for any χ^2 random variable. Then

$$\mathsf{E}\left(h(X_p)\right) \; = \; p\,\mathsf{E}\left(\frac{h(X_{p+2})}{X_{p+2}}\right)$$

Theorem 5.8 (Hwang)

Let $g(x)$ be function such that $g(-1)$ is finite and $\mathsf{E}(g(X))$ exists for the two distributions below then

1. *If $X \sim$ Poisson(λ) then*

$$\mathsf{E}\left(\lambda g(X)\right) \; = \; \mathsf{E}\left(X g(X-1)\right)$$

2. *If $X \sim$ Negative Binomial(r, p) then*

$$\mathsf{E}\left((1-p)g(X)\right) \; = \; \mathsf{E}\left(\frac{X}{r+X-1}\; g(X-1)\right)$$

Some Further Inequalities

Theorem 5.9 (Cauchy-Schwarz Inequality)

Let X and Y be any two random variables then

$$\left|\mathsf{E}(XY)\right| \leqslant \mathsf{E}\left(|XY|\right) \leqslant \sqrt{\mathsf{E}(X^2)\,\mathsf{E}(Y^2)}$$

Corollary 5.9.1

Suppose that X and Y are two random variables with finite means and finite variances then

$$\left|Cov(X,Y)\right| \leqslant \sqrt{\mathsf{Var}(X)\,\mathsf{Var}(Y)}$$

A Generalization of Cauchy-Schwarz

Theorem 5.10 (Hölder's Inequality)

Let X and Y be two random variables and let p and q be positive numbers greater than 1 such that $1/p + 1/q = 1$ then

$$\Big|\mathsf{E}(XY)\Big| \leqslant \mathsf{E}\Big(|XY|\Big) \leqslant \Big(\mathsf{E}\left(|X|^p\right)\Big)^{1/p}\Big(\mathsf{E}\left(|Y|^q\right)\Big)^{1/q}$$

The proof of Hölder's Inequality depends on the following lemma from number theory

Lemma 5.1

Let a and b be any positive numbers and let p and q be positive numbers satisfying $1/p + 1/q = 1$ then

$$\frac{a^p}{p} + \frac{b^q}{q} \geqslant ab$$

Consequences of Hölder's Inequality

Theorem 5.11 (Liapounov's Inequality)

If X is a random variable and $0 < r < s$ then

$$\mathsf{E}\left(|X|^r\right) \;\leqslant\; \mathsf{E}\left(|X|^s\right)^{r/s}$$

Theorem 5.12 (Minkowski's Inequality)

Let X and Y be two random variables and let $p > 1$ then

$$\Big(\mathsf{E}\left(|X+Y|^p\right)\Big)^{1/p} \;\leqslant\; \Big(\mathsf{E}\left(|X|^p\right)\Big)^{1/p} + \Big(\mathsf{E}\left(|Y|^p\right)\Big)^{1/p}$$

Jensen's Inequality

Definition 5.1

A function g is said to be convex if for every x, y and $0 < \lambda < 1$

$$g\Big(\lambda x + (1 - \lambda y)\Big) \;\leqslant\; \lambda g(x) + (1 - \lambda) g(y)$$

A function g is concave if $-g$ is convex and so

$$g\Big(\lambda x + (1 - \lambda y)\Big) \;\geqslant\; \lambda g(x) + (1 - \lambda) g(y)$$

Theorem 5.13 (Jensen's Inequality)

If X is a random variable with finite mean and g is a convex function then

$$\mathsf{E}\Big(g(X)\Big) \;\geqslant\; g\Big(\mathsf{E}(X)\Big)$$