Lecture 11: Inequalities

Mathematical Statistics I, MATH 60061/70061

Tuesday October 12, 2021

Reference: Casella & Berger, 3.6, 4.7

When an analytical solution is not available

Strategies when we can't calculate a probability or expectation exactly:

- Monte Carlo simulations
  - 40024/50024 Computational Statistics
- Bounds using inequalities
  - Bounds on expectations
  - Bounds on tails probabilities
- Approximations using limit theorems
  - The law of large numbers
  - The central limit theorem

Numerical inequality

Lemma: Let a and b be any positive numbers, and let p and q be any positive numbers (necessarily greater than 1) satisfying

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab$$

with equality if and only if $a^p = b^q$.

Proof: Fix b, and consider the function

$$g(a) = \frac{1}{p}a^p + \frac{1}{q}b^q - ab.$$

Then

$$\frac{d}{da}g(a) = 0 \implies a^{p-1} - b = 0 \implies b = a^{p-1}.$$

A check of the second derivative will establish that this is a unique minimum. Since $(p-1)q = p$, the value of the function at minimum is

$$\begin{aligned} \frac{1}{p}a^p + \frac{1}{q}(a^{p-1})^q - aa^{p-1} &= \frac{1}{p}a^p + \frac{1}{q}a^p - a^p \\ &= 0. \end{aligned}$$

The equality holds only if $a^{p-1} = b$, which is equivalent to $a^p = b^q$.

Hölder's inequality

Let X and Y be any two random variables, and let p and q satisfying $p^{-1} + q^{-1} = 1$. Then

$$|E(XY)| \leq E(|XY|) \leq (E|X|^p)^{1/p}(E|Y|^q)^{1/q}.$$

Hölder's inequality

Let X and Y be any two random variables, and let p and q satisfying $p^{-1} + q^{-1} = 1$. Then

$$|E(XY)| \leq E(|XY|) \leq (E|X|^p)^{1/p}(E|Y|^q)^{1/q}.$$

- The first inequality follows because $-|XY| \leq XY \leq |XY|$.
- To prove the second inequality, define

$$a = \frac{|X|}{(E|X|^p)^{1/p}} \text{ and } b = \frac{|Y|}{(E|Y|^q)^{1/q}}.$$

Applying the lemma, we get

$$\frac{1}{p}\frac{|X|^p}{E|X|^p} + \frac{1}{q}\frac{|Y|^q}{E|Y|^q} \geq \frac{|XY|}{(E|X|^p)^{1/p}(E|Y|^q)^{1/q}}.$$

Taking expectations of both sides gives the desired result.

Special cases of Hölder's inequality

- Cauchy-Schwartz inequality ($p = q = 2$): For any two random variables X and Y,

$$|E(XY)| \leq E(|XY|) \leq (E|X|^2)^{1/2}(E|Y|^2)^{1/2}.$$

- If we set $Y \equiv 1$, then

$$E|X| \leq (E|X|^p)^{1/p}$$

for any $p \geq 1$.

Expectation inequalities from Hölder's inequality

- Liapounov's inequality: If r and s are constants satisfying $1 \leq r \leq s$ and X is a random variable, then

$$(E|X|^r)^{1/r} \leq (E|X|^s)^{1/s}.$$

- Minkowski's inequality: If $p \geq 1$ is a constant and X and Y are random variables, then

$$(E|X+Y|^p)^{1/p} \leq (E|X|^p)^{1/p} + (E|Y|^p)^{1/p},$$

  where $(E|X|^p)^{1/p} = \|X\|_p$ defines a length for random variables, and is called L_p norm.

Example: bound on correlation

If X and Y have mean 0, $E(X) = E(Y) = 0$, then

$$E(XY) = \mathrm{Cov}(X, Y),\ E(X^2) = \mathrm{Var}(X),\ E(Y^2) = \mathrm{Var}(Y).$$

By Cauchy-Schwarz,

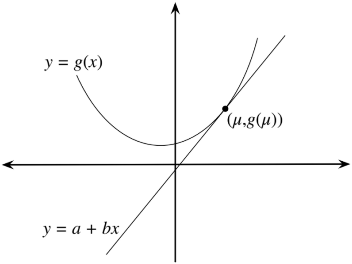
$$\mathrm{Cov}(X, Y)^2 \le \mathrm{Var}(X)\mathrm{Var}(Y) \Rightarrow |\mathrm{Corr}(X, Y)| \le 1.$$

When the means are *not* 0, applying Cauchy-Schwarz to the *centered* random variables $X - E(X)$ and $Y - E(Y)$ gives $|\mathrm{Corr}(X, Y)| \le 1$.

Jensen: an inequality for convexity

Let X be a random variable.

- If g is a convex function, then $E(g(X)) \geq g(E(X))$.
- If g is a concave function, then $E(g(X)) \leq g(E(X))$.



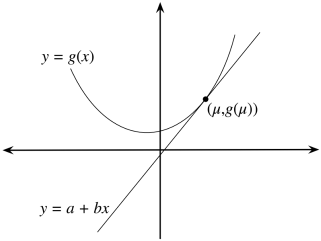
A function g whose domain is an interval I is *convex* if

$$g(px_1+(1-p)x_2) \leq pg(x_1)+(1-p)g(x_2)$$

for all $x_1, x_2 \in I$ and $p \in (0,1)$.

A function g is *concave* if $-g$ is convex.

Jensen: an inequality for convexity, continued



A function g whose domain is an interval I is *convex* if

$$g(px_1+(1-p)x_2) \leq pg(x_1)+(1-p)g(x_2)$$

for all $x_1, x_2 \in I$ and $p \in (0,1)$.

A function g is *concave* if $-g$ is convex.

If g is convex, then all lines tangent to g lie below g. Let $\mu = E(X)$ and consider the tangent point $(\mu, g(\mu))$. Denoting the tangent line by $a + bx$, we have $g(x) \geq a + bx$ for all x, so $g(X) \geq a + bX$. Taking the expectation of both sides:

$$E(g(X)) \geq a + bE(X) = a + b\mu = g(\mu) = g(E(X)).$$

If g is concave, then $h = -g$ is convex. The inequality for g is reversed for the concave case.

Example: inequality for means

If $a_1, \ldots, a_n$ are positive numbers, define

$$a_A = \frac{1}{n}(a_1 + a_2 + \ldots a_n) \qquad \text{[arithmetic mean]}$$

$$a_G = [a_1 a_2 \cdot \cdots \cdot a_n]^{1/n} \qquad \text{[geometric mean]}$$

$$a_H = \frac{1}{\frac{1}{n}\left(\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}\right)} \qquad \text{[harmonic mean]}$$

Using Jensen's inequality, find the relationship between arithmetic mean, geometric mean, and harmonic mean.

Markov's inequality

For any random variable X and constant $r > 0$,

$$P(|X| \geq r) \leq \frac{E|X|}{r}.$$

Markov's inequality

For any random variable X and constant $r > 0$,

$$P(|X| \geq r) \leq \frac{E|X|}{r}.$$

Let $Y = \frac{|X|}{r}$. Then $I(\{Y \geq 1\}) \leq Y$.

- If $I(\{Y \geq 1\}) = 1$, then $Y \geq 1$.
- If $I(\{Y \geq 1\}) = 0$, then $0 \leq Y < 1$.

Taking the expectation of both sides gives Markov's inequality.

Markov's inequality

For any random variable X and constant $r > 0$,

$$P(|X| \geq r) \leq \frac{E|X|}{r}.$$

Let $Y = \frac{|X|}{r}$. Then $I(\{Y \geq 1\}) \leq Y$.

- If $I(\{Y \geq 1\}) = 1$, then $Y \geq 1$.
- If $I(\{Y \geq 1\}) = 0$, then $0 \leq Y < 1$.

Taking the expectation of both sides gives Markov's inequality.

Let X be the income of a randomly selected individual from a population.

- Taking $r = 2E(X)$, $P(X \geq 2E(X)) \leq 1/2$.
- Similarly, $P(X \geq 3E(X)) \leq 1/3$.

Chebyshev's inequality

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

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

Chebyshev's inequality

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

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

Proof:

$$\begin{aligned} E(g(X)) &= \int_{-\infty}^{\infty} g(x) f_X(x) dx \\ &\geq \int_{-\infty}^{\infty} I(\{g(x) \geq r\}) g(x) f_X(x) dx \\ &\geq r \int_{-\infty}^{\infty} I(\{g(x) \geq r\}) f_X(x) dx = r \int_{\{g(x) \geq r\}} f_X(x) dx \\ &= r P(g(X) \geq r) \end{aligned}$$

where $I(A)$ is the indicator function of the set A.

Different forms of Chebychev's inequality

- If g is nondecreasing, then another form of Chebychev's inequality is, for $\epsilon > 0$,

$$P(X \geq \epsilon) \leq \frac{E(g(X))}{g(\epsilon)}$$

- Suppose that X has expectation μ and variance σ^2. For $g(x) = (x-\mu)^2/\sigma^2$, we have

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

- If X has a finite kth moment with an integer k, then, for $t > 0$,

$$P(|X-\mu| \geq t) \leq \frac{E|X-\mu|^k}{t^k}$$

Chebychev's inequality is usually quite conservative:

$$P\left(\frac{|X-\mu|}{\sigma} \geq 1\right) \leq 1$$
$$P\left(\frac{|X-\mu|}{\sigma} \geq 2\right) \leq 1/4$$
$$P\left(\frac{|X-\mu|}{\sigma} \geq 3\right) \leq 1/9$$

Chebychev inequality as a theoretical tool

The Chebychev inequality is useful as a theoretical tool. Examples:

- Suppose $g(X) \geq 0$ and $E(g(X)) = 0$, then

$$P(g(X) \geq r) = 0$$

  for all $r > 0$, i.e., $P(g(X) = 0) = 1$.
- If $\mathrm{Var}(X) = 0$ then $P(X = E(X)) = 1$.

Weak law of large numbers

Suppose $X_1, X_2, \ldots, X_n$ are independent with common mean μ and common variance σ^2. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then

$$\mathrm{Var}(\bar{X}) = \frac{1}{n^2}\mathrm{Var}(\sum_{i=1}^{n} X_i) = \frac{\sigma^2}{n}.$$

So, for any $\epsilon > 0$,

$$P(|\bar{X} - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} \to 0,$$

as $n \to \infty$