Mathematical Statistics 1

Chapter 1. Introduction to Probability

1.10. Important Inequalities—Proofs of Theorems

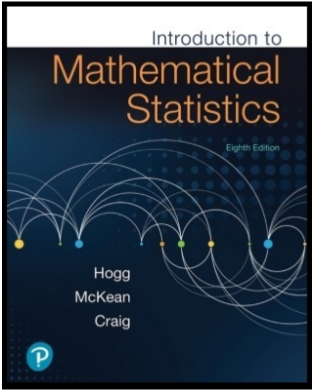

Table of contents

Theorem 1.10.1

Theorem 1.10.1. Let X be a random variable and let$m \in \mathbb{N}$. Suppose $E[|X|^m]$ exists. If $k \in \mathbb{N}$ and $k \leq m$, then $E(X^k)$ exists.

Proof. We give a proof for X a continuous random variable and leave the discrete case as an exercise. Let f be the probability density function for X. We apply Theorem 1.8.1(a) with $g(X) = X^k$. We have

$$\begin{aligned}
\int_{-\infty}^{\infty} |g(x)| f(x)\,dx &= \int_{-\infty}^{\infty} |x|^k f(x)\,dx \\
&= \int_{|x|\leq 1} |x|^k f(x)\,dx + \int_{|x|>1} |x|^m f(x)\,dx \\
&\leq \int_{-\infty}^{\infty} f(x)\,dx + \int_{-\infty}^{\infty} |x|^m f(x)\,dx \text{ since } f(x) \geq 0 \\
&\leq \int_{-\infty}^{\infty} f(x)\,dx + \int_{-\infty}^{\infty} |x|^m f(x)\,dx \\
&\leq 1 + E[|X|^m] < \infty \text{ by hypothesis.}
\end{aligned}$$

Theorem 1.10.1

Theorem 1.10.1. Let X be a random variable and let$m \in \mathbb{N}$. Suppose $E[|X|^m]$ exists. If $k \in \mathbb{N}$ and $k \leq m$, then $E(X^k)$ exists.

Proof. We give a proof for X a continuous random variable and leave the discrete case as an exercise. Let f be the probability density function for X. We apply Theorem 1.8.1(a) with $g(X) = X^k$. We have

$$\begin{aligned}
\int_{-\infty}^{\infty} |g(x)| f(x)\, dx &= \int_{-\infty}^{\infty} |x|^k f(x)\, dx \\
&= \int_{|x| \leq 1} |x|^k f(x)\, dx + \int_{|x|>1} |x|^m f(x)\, dx \\
&\leq \int_{-\infty}^{\infty} f(x)\, dx + \int_{-\infty}^{\infty} |x|^m f(x)\, dx \text{ since } f(x) \geq 0 \\
&\leq \int_{-\infty}^{\infty} f(x)\, dx + \int_{-\infty}^{\infty} |x|^m f(x)\, dx \\
&\leq 1 + E[|X|^m] < \infty \text{ by hypothesis.}
\end{aligned}$$

Theorem 1.10.1 (continued)

Theorem 1.10.1. Let X be a random variable and let$m \in \mathbb{N}$. Suppose $E[|X|^m]$ exists. If $k \in \mathbb{N}$ and $k \leq m$, then $E(X^k)$ exists.

Proof (continued). So by Theorem 1.8.1(a),

$$\int_{-\infty}^{\infty} g(x)f(x)\,dx = \int_{-\infty}^{\infty} x^k f(x)\,dx = E[X^k] < \infty$$

and so $E[X^k]$ exists for all $k \in \mathbb{N}$, $k \leq m$, as claimed. □

Exercise 1.10.2

Theorem 1.10.2. Markov's Inequality.
Let $u(X)$ be a nonnegative function of random variable X. If $E[u(X)]$ exists then for every positive constant c, $P(u(x) \geq c) \leq E[u(X)]/c$.

Proof. We give a proof for X a continuous random variable and leave the discrete case as an exercise. Let $A = \{x \mid u(x) \geq c\}$ and let f denote the probability density function of X. Then

$$E[u(X)] = \int_{-\infty}^{\infty} u(x)f(x) = \int_A u(x)f(x)\,dx + \int_{\mathbb{R}\setminus A} u(x)f(x)\,dx$$

(we need u and f to be measurable and we need Lebesgue integrals here so that we know both of the two integrals on the right are defined). Since each of the two integrals on the right are nonnegative then

$$E[u(X)] \geq \int_A u(x)f(x)\,dx \geq \int_A cf(x)\,dx$$

since $u(x) \geq c$ for $x \in A$.

Exercise 1.10.2

Theorem 1.10.2. Markov's Inequality.
Let $u(X)$ be a nonnegative function of random variable X. If $E[u(X)]$ exists then for every positive constant c, $P(u(x) \geq c) \leq E[u(X)]/c$.

Proof. We give a proof for X a continuous random variable and leave the discrete case as an exercise. Let $A = \{x \mid u(x) \geq c\}$ and let f denote the probability density function of X. Then

$$E[u(X)] = \int_{-\infty}^{\infty} u(x)f(x) = \int_A u(x)f(x)\,dx + \int_{\mathbb{R}\setminus A} u(x)f(x)\,dx$$

(we need u and f to be measurable and we need Lebesgue integrals here so that we know both of the two integrals on the right are defined). Since each of the two integrals on the right are nonnegative then

$$E[u(X)] \geq \int_A u(x)f(x)\,dx \geq \int_A cf(x)\,dx$$

since $u(x) \geq c$ for $x \in A$.

Exercise 1.10.2 (continued)

Theorem 1.10.2. Markov's Inequality.
Let $u(X)$ be a nonnegative function of random variable X. If $E[u(X)]$ exists then for every positive constant c, $P(u(x) \geq c) \leq E[u(X)]/c$.

Proof (continued). Now

$$P(X \in A) = P(u(x) \geq c) = \int_A f(x)\,dx \leq \frac{E[u(X)]}{c},$$

as claimed. □

Theorem 1.10.3

Theorem 1.10.3. Chebyshev's Inequality.
Let X be a random variable where $E(X^2) < \infty$ (so that μ and σ^2 are define). Then for every $k > 0$

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \text{ or } P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

Proof. Define $u(X) = (X - \mu)^2$ and $c = k^2\sigma^2$. Then by Theorem 1.10.2, Markiv's Inequality,

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

That is $P(|X - \mu| \geq k\sigma) \leq 1/k^2$, as claimed. □

Theorem 1.10.3

Theorem 1.10.3. Chebyshev's Inequality.
Let X be a random variable where $E(X^2) < \infty$ (so that μ and σ^2 are define). Then for every $k > 0$

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \text{ or } P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

Proof. Define $u(X) = (X - \mu)^2$ and $c = k^2\sigma^2$. Then by Theorem 1.10.2, Markiv's Inequality,

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

That is $P(|X - \mu| \geq k\sigma) \leq 1/k^2$, as claimed. □

Theorem 1.10.5

Theorem 1.10.5. Jensen's Inequality.
If φ is convex on an open interval I and X is a random variable whose support is contained in I and has finite expectation, then $\varphi(E[X]) \leq E[\varphi(X)]$. If φ is strictly convex, then the inequality is strict unless X is a constant random variable.

Proof. As stated above, we give a proof assuming that φ is twice differentiable. Then by Taylor's Theorem (see my online Calculus 2 [MATH 1920] notes on 10.9. Convergence of Taylor Series, Theorem 23) we have for any $u \in I$ that

$$\varphi(x) = \varphi(u) + \varphi'(u)(x - \mu) + \frac{\varphi''(\zeta)(x - \mu)^2}{2!}$$

for a given $x \in I$ and for some ζ between x and μ. Since $\varphi''(\zeta) \geq 0$ by Theorem 1.10.4 then $\varphi''(\zeta)(x - \mu)^2/2 \geq 0$ and hence $\varphi(x) \geq \varphi(u) + \varphi'(u)(x - \mu)$.

Theorem 1.10.5

Theorem 1.10.5. Jensen's Inequality.
If φ is convex on an open interval I and X is a random variable whose support is contained in I and has finite expectation, then $\varphi(E[X]) \leq E[\varphi(X)]$. If φ is strictly convex, then the inequality is strict unless X is a constant random variable.

Proof. As stated above, we give a proof assuming that φ is twice differentiable. Then by Taylor's Theorem (see my online Calculus 2 [MATH 1920] notes on 10.9. Convergence of Taylor Series, Theorem 23) we have for any $u \in I$ that

$$\varphi(x) = \varphi(u) + \varphi'(u)(x - \mu) + \frac{\varphi''(\zeta)(x - \mu)^2}{2!}$$

for a given $x \in I$ and for some ζ between x and μ. Since $\varphi''(\zeta) \geq 0$ by Theorem 1.10.4 then $\varphi''(\zeta)(x - \mu)^2/2 \geq 0$ and hence $\varphi(x) \geq \varphi(u) + \varphi'(u)(x - \mu)$.

Theorem 1.10.5 (continued)

Theorem 1.10.5. Jensen's Inequality.
If φ is convex on an open interval I and X is a random variable whose support is contained in I and has finite expectation, then $\varphi(E[X]) \leq E[\varphi(X)]$. If φ is strictly convex, then the inequality is strict unless X is a constant random variable.

Proof (continued). Therefore

$$\int_X \varphi(x) f(x)\,dx \geq \int_X \varphi(u) f(x)\,dx + \varphi'(u) \int_X (x-\mu) f(x)\,dx$$

or

$$E[\varphi(X)] \geq \varphi(u) \int_X f(x)\,dx + \varphi'(u) E[X-\mu] = \varphi(u).$$

Since this holds for any $x, u \in I$, then we take $u = E[X] \in I$ to et $\varphi(E[X]) \leq E[\varphi(X)]$, as claimed. The inequality is strict if $\varphi''(x) > 0$ for all $x \in I$, provided X is not a constant. □