Moments and moment generating function

Dr. Qiuzhuang Sun

STAT3023

Review of random variables (STAT2011/2911)

Consider a sample space Ω with a probability measure P. Let X be a random variable defined on this sample space.

- Any random variable X has a cumulative distribution function (cdf), $F_X(x) = P(X \leq x)$.
- Discrete random variables:
  - X gets realized values on a countable set
  - Probability mass function (pmf): $P(X = x) = p_X(x)$
- Continuous random variables:
  - X gets realized values on an uncountable set
  - Example: $X \sim \text{Unif}(0, 1)$
  - Probability density function (pdf): $f_X(x) = \frac{dF_X(x)}{dx}$
  - $f_X(x) \geq 0$; $\int_{-\infty}^{\infty} f_X(x)dx = 1$
  - $P(a \leq X \leq b) = \int_a^b f_X(x)dx$

Moments

For any random variable X and a function $g : \mathbb{R} \to \mathbb{R}$, define the **expectation** of $g(X)$ to be

$$E[g(X)] = \begin{cases} \sum_x g(x) p_X(x), & X \text{ is discrete,} \\ \int_{-\infty}^{\infty} g(x) f_X(x) dx, & X \text{ is continuous.} \end{cases}$$

(provided the sum or the integral is finite)

Examples:

- rth moment: $\mu_r = E(X^r)$
  - $\mu = \mu_1 = E(X)$
- rth central moment: $E[(X - \mu)^r]$
  - $\text{var}(X) = E(X^2) - \mu^2 = \mu_2 - \mu^2$

Moment generating functions

- Moment generating function: encoding the sequence of moments $\{E(X^r)\}$, $r = 1, 2, \ldots, \infty$ into the coefficients of a power series.

- Choose $g(x) = \exp(tx)$, then the moment generating function (mgf) of a random variable X is defined to be

$$M_X(t) = E\{g(X)\} = E\{\exp(tX)\},$$

provided this expectation exists for t in some open interval containing zero.

Getting moments from mgf

Examples

- $X \sim \mathsf{Unif}(0,1)$

Examples

- $X \sim \text{Bern}(p)$

- $X \sim \text{Binomial}(n, p)$

Examples

- $X \sim N(0, 1)$

Uniqueness of mgf

- If the moment generating functions exists, and $M_X(t) = M_Y(t)$, then X and Y have the same distributions.

- Nevertheless, if two random variables have all the same moments, $E(X^r) = E(Y^r)$ for all $r = 1, 2, \ldots$ then X and Y do not necessarily have the same distributions.

Example: X is standard log-normal with pdf

$$f_X(x) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}[\log(x)]^2}, \quad x > 0,$$

and Y has pdf

$$f_Y(y) = f_X(y)[1 + \sin(2\pi \log(y))], \quad y > 0.$$

Showing $E[X^r] = E[Y^r]$

$$f_X(x) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}[\log(x)]^2}, \quad x > 0;$$

$$f_Y(y) = f_X(y)[1 + \sin(2\pi \log(y))], \quad y > 0.$$

Properties of mgf

Let X be a RV with mgf $M_X(t)$. Then the random variable $Z = aX + b$ has mgf $M_Z(t) = e^{tb} M_X(at)$.

- Derive $M_Z(t)$ for $Z \sim N(\mu, \sigma^2)$.

Properties of mgf

Recall for two **independent** random variables X and Y, we have

$$E\left\{g(X)h(Y)\right\} = E\left\{g(X)\right\} E\left\{h(Y)\right\}$$

for any two functions g and h. Let $M_X(t)$ and $M_Y(t)$ be mgfs of X and Y respectively, then the mgf of $Z = X + Y$ is given by $M_Z(t) = M_X(t)M_Y(t)$.

Sum of independent random variables

More generally, if $X_1, \ldots, X_n$ be mutually independent random variables with mgfs $M_{X_i}(t)$ for $i = 1, \ldots, n$, then the mgf of $Z = \sum_{i=1}^{n} X_i$ is given by $M_Z(t) = \prod_{i=1}^{n} M_{X_i}(t)$.

Example: Let $X_1, \ldots, X_n$ be independent and $X_i \sim N(\mu_i, \sigma_i^2)$. What is the distribution of $Z = \sum_{i=1}^{n} X_i$?

Examples

Example: Let $X_1, \ldots, X_n$ be independent and $X_i \sim \text{Poisson}(\lambda_i)$. What is the distribution of $Z = \sum_{i=1}^{n} X_i$?

Probability bounds

Markov's inequality: For any non-negative random variable X and any $a > 0$, we have

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

Proof:

Probability bounds

Chebyshev's inequality: For any random variable X and any $a > 0$, we have

$$P\left(|X - E(X)| \geq a\right) \leq \frac{\mathsf{Var}(X)}{a^2}.$$

Proof:

Probability bounds

Chernoff's bounds: For any random variable X, we have

$$P(X \geq a) = P(e^{tX} \geq e^{ta}) \leq e^{-ta} M_X(t), \quad t > 0.$$

This implies

$$P(X \geq a) \leq \inf_{t>0} e^{-ta} M_X(t).$$

Examples

Let $X \sim \text{Binomial}(n, p)$. Derive bounds for $P(X \geq \alpha n)$ for $p < \alpha < 1$.

Examples

Let $X \sim \text{Binomial}(n, p)$. Derive bounds for $P(X \geq \alpha n)$ for $p < \alpha < 1$.

Convergence of mgfs implies convergences of cdfs

Suppose $X_1, X_2, \ldots,$ is a sequence of random variables, each with mgf $M_{X_n}(t)$. Furthermore, suppose that

$$\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$$

for all t in an open interval containing zero, and $M_X(t)$ is the mgf of a random variable X. Then for any x such that $F_X(x)$ is continuous, we have

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x).$$

We also say that the sequence $X_1, X_2, \ldots, X_n$ converges to X in distribution.

Application: Poisson approximation to binomial distribution

Let $X_n \sim \text{Binomial}(n, p)$. When $n \to \infty$, we assume $np \to \lambda > 0$. Then $X_n \xrightarrow{d} X$ with $X \sim \text{Poisson}(\lambda)$.

Application: Central limit theorem

Let $X_1, \ldots, X_n$ be i.i.d. with $E[X] = \mu$ and var$(X) = \sigma^2$. Assume the mgf for X_i exists. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. We have $Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$.

Application: Central limit theorem

Application: Central limit theorem

Convergence in probability

The sequence of random variables $X_1, \ldots, X_n$ is said to converge to a random variable X in probability if $\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0$ for any $\epsilon > 0$.

We write $X_n \xrightarrow{P} X$ if X_n converges to X in probability. We write $X_n \xrightarrow{P} a$ if X_n converges to a constant a in probability.

Weak law of large numbers

Let $X_1, \ldots, X_n$ be i.i.d. with $E[X_i] = \mu$ and $\mathrm{var}(X_i) = \sigma^2$. Let $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Then $\bar{X}_n \xrightarrow{P} \mu$.