Advanced statistics and modelling

2020. február 25.

Expectation, Inequalities, Convergence

Covariance and correlation

Covariance and correlation

The covariance between random variables X and Y is defined as

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))].$$

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}$$

$$\mathbb{V}\left(\sum_{i}a_{i}X_{i}\right)=\sum_{i}a_{i}^{2}\mathbb{V}(X_{i})+2\sum_{i\leq i}a_{i}a_{i}\mathsf{Cov}(X_{i},X_{j}).$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation

Covariance matrix Conditional expectation Characteristic

Inequalities

Markov's inequality
Chebyshev's
inequality

Cauchy–Schwarz inequality

Jensen's inequality

Convergence Convergence types Law of large

numbers Central limit

Covariance and correlation

The covariance between random variables X and Y is defined as

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))].$$

• The **correlation** is simply

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}$$

 According to the definition of the covariance Cov(X, Y) = E(XY) − E(X)E(Y).

The correlation satisfies $-1 \le Corr(X, Y) \le 1$

For random variables $X_1, X_2, ..., X_t$ the variance of their linear combination is

$$\mathbb{V}\left(\sum a_i X_i\right) = \sum a_i^2 \mathbb{V}(X_i) + 2 \sum_i a_i a_j \text{Cov}(X_i, X_j).$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation

Covariance matrix Conditional expectation Characteristic

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz

Convergence Convergence types

Law of large numbers

Central limit

Covariance and correlation

• The **covariance** between random variables *X* and *Y* is defined as

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))].$$

The correlation is simply

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}$$

- According to the definition of the covariance $Cov(X, Y) = \mathbb{E}(XY) \mathbb{E}(X)\mathbb{E}(Y)$.
- The correlation satisfies $-1 \le Corr(X, Y) \le 1$.
- For random variables $X_1, X_2, \dots X_i$ the variance of their linear combination is

$$\mathbb{V}\left(\sum_{i} a_{i} X_{i}\right) = \sum_{i} a_{i}^{2} \mathbb{V}(X_{i}) + 2 \sum_{i < j} a_{i} a_{j} \operatorname{Cov}(X_{i}, X_{j}).$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation

Covariance matri Conditional expectation Characteristic

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality

Convergence
Convergence types
Law of large

numbers Central limit

Covariance and correlation

• The **covariance** between random variables *X* and *Y* is defined as

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))].$$

The correlation is simply

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}$$

- According to the definition of the covariance Cov(X, Y) = E(XY) - E(X)E(Y).
- The correlation satisfies $-1 \le Corr(X, Y) \le 1$.
- For random variables $X_1, X_2, ... X_i$ the variance of their linear combination is

$$\mathbb{V}\left(\sum_{i} a_{i} X_{i}\right) = \sum_{i} a_{i}^{2} \mathbb{V}(X_{i}) + 2 \sum_{i < j} a_{i} a_{j} \operatorname{Cov}(X_{i}, X_{j}).$$

Expectation, Inequalities, Convergence

Expectation and Variance Covariance and

Covariance and correlation

Covariance mater Conditional expectation Characteristic function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality

Convergence Convergence type Law of large

Law of large numbers Central limi

Covariance and correlation

• The **covariance** between random variables *X* and *Y* is defined as

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))].$$

The correlation is simply

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}$$

- According to the definition of the covariance $Cov(X, Y) = \mathbb{E}(XY) \mathbb{E}(X)\mathbb{E}(Y)$.
- The correlation satisfies $-1 \le Corr(X, Y) \le 1$.
- For random variables $X_1, X_2, \dots X_i$ the variance of their linear combination is

$$\mathbb{V}\left(\sum_{i} a_{i} X_{i}\right) = \sum_{i} a_{i}^{2} \mathbb{V}(X_{i}) + 2 \sum_{i < j} a_{i} a_{j} \operatorname{Cov}(X_{i}, X_{j}).$$

.

Expectation, Inequalities, Convergence

Expectation and Variance Covariance and

correlation

Covariance matrix
Conditional
expectation
Characteristic

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence Convergence types Law of large

numbers Central limit

Variance-covariance matrix

For a vector of random variables $X = (X_1, X_2, ..., X_n)$ the **variance-covariance matrix** is defined as

$$\mathbb{V}(X) = \left(\begin{array}{cccc} \mathbb{V}(X_1) & \operatorname{Cov}(X_1, X_2) & \cdots & \operatorname{Cov}(X_1, X_n) \\ \operatorname{Cov}(X_2, X_1) & \mathbb{V}(X_2) & \cdots & \operatorname{Cov}(X_2, X_n) \\ \vdots & & \ddots & \vdots \\ \operatorname{Cov}(X_n, X_1) & \operatorname{Cov}(X_n, X_2) & \cdots & \mathbb{V}(X_n) \end{array} \right)$$

 For a multinomial distribution X ~Multinomial(N,p) the variance-covariance matrix reads

$$\mathbb{V}(X) = \begin{pmatrix} Np_{1}(1-p_{1}) & -Np_{1}p_{2} & \cdots & -Np_{1}p_{n} \\ -Np_{2}p_{1} & Np_{2}(1-p_{2}) & \cdots & -Np_{2}p_{n} \\ \vdots & & \ddots & \vdots \\ -Np_{k}p_{1} & -Np_{k}p_{2} & \cdots & Np_{k}(1-p_{k}) \end{pmatrix}$$

• In case of a multivariate Normal distribution, $X \sim N(\mu, \Sigma)$ the parameter Σ is actually the covariance matrix.

Expectation, Inequalities, Convergence

Variance
Covariance and
correlation
Covariance matrix
Conditional

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequalit Cauchy-Schwarz inequality Jensen's inequality

Convergence Convergence types Law of large numbers

Variance-covariance matrix

For a vector of random variables $X = (X_1, X_2, \dots, X_n)$ the **variance-covariance matrix** is defined as

$$\mathbb{V}(X) = \left(\begin{array}{cccc} \mathbb{V}(X_1) & \operatorname{Cov}(X_1, X_2) & \cdots & \operatorname{Cov}(X_1, X_n) \\ \operatorname{Cov}(X_2, X_1) & \mathbb{V}(X_2) & \cdots & \operatorname{Cov}(X_2, X_n) \\ \vdots & & \ddots & \vdots \\ \operatorname{Cov}(X_n, X_1) & \operatorname{Cov}(X_n, X_2) & \cdots & \mathbb{V}(X_n) \end{array} \right)$$

 For a multinomial distribution X ~Multinomial(N,p) the variance-covariance matrix reads

$$\mathbb{V}(X) = \begin{pmatrix} Np_1(1-p_1) & -Np_1p_2 & \cdots & -Np_1p_n \\ -Np_2p_1 & Np_2(1-p_2) & \cdots & -Np_2p_n \\ \vdots & & \ddots & \vdots \\ -Np_kp_1 & -Np_kp_2 & \cdots & Np_k(1-p_k) \end{pmatrix}$$

• In case of a multivariate Normal distribution, $X \sim N(\mu, \Sigma)$ the parameter Σ is actually the covariance matrix.

ĸ.

Expectation, Inequalities, Convergence

Variance
Covariance and correlation
Covariance matrix
Conditional expectation
Characteristic

Inequalitie

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence
Convergence types
Law of large
numbers

Variance-covariance matrix

For a vector of random variables $X = (X_1, X_2, ..., X_n)$ the variance-covariance matrix is defined as

$$\mathbb{V}(X) = \begin{pmatrix} \mathbb{V}(X_1) & \operatorname{Cov}(X_1, X_2) & \cdots & \operatorname{Cov}(X_1, X_n) \\ \operatorname{Cov}(X_2, X_1) & \mathbb{V}(X_2) & \cdots & \operatorname{Cov}(X_2, X_n) \\ \vdots & \ddots & \vdots \\ \operatorname{Cov}(X_n, X_1) & \operatorname{Cov}(X_n, X_2) & \cdots & \mathbb{V}(X_n) \end{pmatrix}$$

 For a multinomial distribution X ~Multinomial(N,p) the variance-covariance matrix reads

$$\mathbb{V}(X) = \begin{pmatrix} Np_1(1-p_1) & -Np_1p_2 & \cdots & -Np_1p_n \\ -Np_2p_1 & Np_2(1-p_2) & \cdots & -Np_2p_n \\ \vdots & & \ddots & \vdots \\ -Np_kp_1 & -Np_kp_2 & \cdots & Np_k(1-p_k) \end{pmatrix}$$

 In case of a multivariate Normal distribution, X ~ N(μ, Σ) the parameter Σ is actually the covariance matrix.

7

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance matrix
Conditional
expectation
Characteristic
function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality

Jensen's inequality

Convergence

Convergence types

numbers
Central limit

• If $\vec{a}=(a_1,a_2,\ldots,a_n)$ is a constant vector, and $\vec{X}=(X_1,X_2,\ldots,X_n)$ is a random vector with $\vec{\mu}=(\mu_1,\mu_2,\ldots,\mu_n)=(\mathbb{E}(X_1),\mathbb{E}(X_2),\ldots,\mathbb{E}(X_n))$ and a variance-covariance matrix $\Sigma_{ij}=\operatorname{Cov}(X_i,X_j)$, then

$$\mathbb{V}(\vec{a}^T\vec{X}) = \vec{a}^T \Sigma \vec{a}.$$

• If A is an n by n matrix, then

$$\mathbb{V}(A\vec{X}) = A\Sigma A^T$$

Expectation, Inequalities, Convergence

Expectation and Variance

correlation

Covariance matrix

Conditional
expectation

Characteristic
function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality

Convergence
Convergence types
Law of large
numbers

• If $\vec{a}=(a_1,a_2,\ldots,a_n)$ is a constant vector, and $\vec{X}=(X_1,X_2,\ldots,X_n)$ is a random vector with $\vec{\mu}=(\mu_1,\mu_2,\ldots,\mu_n)=(\mathbb{E}(X_1),\mathbb{E}(X_2),\ldots,\mathbb{E}(X_n))$ and a variance-covariance matrix $\Sigma_{ij}=\operatorname{Cov}(X_i,X_j)$, then

$$\mathbb{V}(\vec{a}^T\vec{X}) = \vec{a}^T \Sigma \vec{a}.$$

• If A is an n by n matrix, then

$$\mathbb{V}(A\vec{X}) = A\Sigma A^T.$$

Expectation, Inequalities, Convergence

Expectation and Variance

correlation

Conditional expectation

Characteris

nequalities

Chebyshev's inequality
Hoeffding's inequality
Cauchy-Schwarz

Convergence types

numbers
Central limit

Conditional expectation

 The conditional expectation and variance of variable X given that variable Y takes the value Y = y is given by

$$\mathbb{E}(X \mid Y = y) = \int x \rho(x \mid Y = y) dx,$$

$$\mathbb{V}(X \mid Y = y) = \int [x - \mathbb{E}(X \mid Y = y)]^2 \rho(x \mid Y = y) dx.$$

(Note that these are functions of y, thus, can be treated also as random variables).

• The conditional expectation of a function r(x, y) is similarly

$$\mathbb{E}(r(X,Y)\mid Y=y) = \int r(x,y)\rho(x\mid Y=y)dx.$$

Rule of iterated expectations

$$\mathbb{E}\left(\mathbb{E}(X\mid Y=y)\right) = \mathbb{E}(X).$$

$$\mathbb{E}\left(\mathbb{E}(r(X,Y)\mid Y=y)\right) = \mathbb{E}(r(X,Y)).$$

Expectation, Inequalities, Convergence

Expectation and Variance

correlation

Conditional expectation

Characteristi function

nequalitie

Markov's inequality
Chebyshev's
inequality
Hoeffding's inequality
Cauchy-Schwarz
inequality
Jensen's inequality

Convergence type Law of large numbers Central limit

Conditional expectation

 The conditional expectation and variance of variable X given that variable Y takes the value Y = y is given by

$$\mathbb{E}(X \mid Y = y) = \int x \rho(x \mid Y = y) dx,$$

$$\mathbb{V}(X \mid Y = y) = \int [x - \mathbb{E}(X \mid Y = y)]^2 \rho(x \mid Y = y) dx.$$

(Note that these are functions of y, thus, can be treated also as random variables).

• The conditional expectation of a function r(x, y) is similarly

$$\mathbb{E}(r(X,Y)\mid Y=y) = \int r(x,y)\rho(x\mid Y=y)dx.$$

Rule of iterated expectations

$$\mathbb{E}\left(\mathbb{E}(X\mid Y=y)\right) = \mathbb{E}(X).$$

$$\mathbb{E}\left(\mathbb{E}(r(X,Y)\mid Y=y)\right) = \mathbb{E}(r(X,Y))$$

Expectation, Inequalities, Convergence

Conditional expectation

Conditional expectation

 The conditional expectation and variance of variable X given that variable Y takes the value Y = v is given by

$$\mathbb{E}(X \mid Y = y) = \int x \rho(x \mid Y = y) dx,$$

$$\mathbb{V}(X \mid Y = y) = \int [x - \mathbb{E}(X \mid Y = y)]^2 \rho(x \mid Y = y) dx.$$

(Note that these are functions of y, thus, can be treated also as random variables).

• The conditional expectation of a function r(x, y) is similarly

$$\mathbb{E}(r(X,Y)\mid Y=y)=\int r(x,y)\rho(x\mid Y=y)dx.$$

$$\mathbb{E}(\mathbb{E}(X \mid Y = y)) = \mathbb{E}(X).$$

$$\mathbb{E}(\mathbb{E}(r(X, Y) \mid Y = y)) = \mathbb{E}(r(X, Y)).$$

Expectation, Inequalities, Convergence

Expectation an Variance

Covariance and correlation

Conditional expectation

Characteristic

Inoqualitio

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence type: Law of large numbers Central limit

Conditional expectation

 The conditional expectation and variance of variable X given that variable Y takes the value Y = y is given by

$$\mathbb{E}(X \mid Y = y) = \int x \rho(x \mid Y = y) dx,$$

$$\mathbb{V}(X \mid Y = y) = \int [x - \mathbb{E}(X \mid Y = y)]^2 \rho(x \mid Y = y) dx.$$

(Note that these are functions of y, thus, can be treated also as random variables).

• The conditional expectation of a function r(x, y) is similarly

$$\mathbb{E}(r(X,Y)\mid Y=y)=\int r(x,y)\rho(x\mid Y=y)dx.$$

· Rule of iterated expectations:

$$\mathbb{E}(\mathbb{E}(X \mid Y = y)) = \mathbb{E}(X).$$

$$\mathbb{E}(\mathbb{E}(r(X, Y) \mid Y = y)) = \mathbb{E}(r(X, Y)).$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance an correlation

Covariance m

Conditional expectation

Character

function

mequalitie

Chebyshev's

inequality
Hoeffding's inequal

inequality

Convergence

Law of large

Central limi

Variance and conditional variance:

$$\mathbb{V}(X) = \mathbb{E}[\mathbb{V}(X \mid Y = y)] + \mathbb{V}[\mathbb{E}(X \mid Y = y)].$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation

Covariance ma

Conditional

Characteristic function

Inequalities

Markov's inequality

Chebyshev's inequality

Cauchy-Schwarz inequality

Convergence type

numbers Central limit

Central limit theorems

Characteristic function and moment generating function

• The characteristic function ϕ of a random variable X is defined as

$$\phi_X(t) = \mathbb{E}\left(e^{itX}\right) = \begin{cases} & \sum_i e^{itx_i} f_X(x_i) & \text{if } X \text{ is discrete,} \\ \\ & \int e^{itx} \rho_X(x) dx & \text{if } X \text{ is continuos.} \end{cases}$$

• The moment generating function ψ is

$$\psi_X(t) = \mathbb{E}\left(e^{tX}\right) = \begin{cases} \sum_i e^{tx_i} f_X(x_i) & \text{if } X \text{ is discrete,} \\ \\ \int e^{tx} \rho_X(x) dx & \text{if } X \text{ is continuo} \end{cases}$$

The n-th moment of X can be calculated as

$$\mathbb{E}(X^n) = \frac{1}{i^n} \frac{\partial^n \phi(t)}{\partial t^n} \bigg|_{t=0} = \frac{\partial^n \psi_X(t)}{\partial t^n} \bigg|_{t=0}$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation Covariance matri

Conditional expectation

Characteristic function

Inequalities

Chebyshev's inequality

Cauchy–Schwarz inequality
Jensen's inequality

Convergence type:

numbers
Central limit

Characteristic function and moment generating function

• The characteristic function ϕ of a random variable X is defined as

$$\phi_X(t) = \mathbb{E}\left(e^{itX}\right) = \left\{ \begin{array}{ll} \sum_i e^{itx_i} f_X(x_i) & \text{if } X \text{ is discrete,} \\ \\ \int e^{itx} \rho_X(x) dx & \text{if } X \text{ is continuos.} \end{array} \right.$$

• The moment generating function ψ is

$$\psi_X(t) = \mathbb{E}\left(e^{tX}\right) = \left\{ \begin{array}{ll} \sum_i e^{tx_i} f_X(x_i) & \text{if } X \text{ is discrete,} \\ \\ \int e^{tx} \rho_X(x) dx & \text{if } X \text{ is continuos.} \end{array} \right.$$

The n-th moment of X can be calculated as

$$\mathbb{E}(X^n) = \frac{1}{i^n} \frac{\partial^n \phi(t)}{\partial t^n} \bigg|_{t=0} = \frac{\partial^n \psi_X(t)}{\partial t^n} \bigg|_{t=0}$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation
Covariance matri
Conditional

Characteristic function

Inequalities

Chebyshev's inequality Hoeffding's inequalit Cauchy-Schwarz inequality Jensen's inequality

Convergence Convergence types Law of large

numbers
Central limit theorems

Characteristic function and moment generating function

• The characteristic function ϕ of a random variable X is defined as

$$\phi_X(t) = \mathbb{E}\left(e^{itX}\right) = \left\{ \begin{array}{ll} \sum_i e^{itx_i} f_X(x_i) & \text{if } X \text{ is discrete,} \\ \\ \int e^{itx} \rho_X(x) dx & \text{if } X \text{ is continuos.} \end{array} \right.$$

• The moment generating function ψ is

$$\psi_X(t) = \mathbb{E}\left(e^{tX}\right) = \left\{ \begin{array}{ll} \sum_i e^{tx_i} f_X(x_i) & \text{if } X \text{ is discrete,} \\ \\ \int e^{tx} \rho_X(x) dx & \text{if } X \text{ is continuos.} \end{array} \right.$$

The n-th moment of X can be calculated as

$$\mathbb{E}(X^n) = \frac{1}{i^n} \frac{\partial^n \phi(t)}{\partial t^n} \bigg|_{t=0} = \frac{\partial^n \psi_X(t)}{\partial t^n} \bigg|_{t=0}.$$

Convergence
Convergence types
Law of large

Central limit

For discrete variables the generating function is defined as

$$G_X(z) = \sum_i z^{x_i} f_X(x_i).$$

- → The characteristic function is $\phi_X(t) = G_X(z = e^u)$ and the moment generating function is $\psi_X(t) = G_X(z = e^t)$.
- For independent random variables X_1, X_2, \ldots, X_n , the characteristic function, the moment generating function (and for discrete variables the generating function) for their sum $X = \sum_{i=1}^{n} X_i$ is simply

$$\phi_X(t) = \prod_{i=1}^n \phi_{X_i}(t), \qquad \psi_X(t) = \prod_{i=1}^n \psi_{X_i}(t), \qquad G_X(z) = \prod_{i=1}^n G_{X_i}(z)$$

Expectation, Inequalities, Convergence

Expectation an Variance

correlation
Covariance ma

Conditional expectation Characteristic

function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence
Convergence types
Law of large
numbers

Central limit

For discrete variables the generating function is defined as

$$G_X(z) = \sum_i z^{x_i} f_X(x_i).$$

- \rightarrow The characteristic function is $\phi_X(t) = G_X(z = e^{it})$ and the moment generating function is $\psi_X(t) = G_X(z = e^t)$.
 - For independent random variables X_1, X_2, \ldots, X_n , the characteristic function, the moment generating function (and for discrete variables the generating function) for their sum $X = \sum_{i=1}^{n} X_i$ is simply

$$\phi_X(t) = \prod_{i=1}^n \phi_{X_i}(t), \qquad \psi_X(t) = \prod_{i=1}^n \psi_{X_i}(t), \qquad G_X(z) = \prod_{i=1}^n G_{X_i}(z)$$

Expectation, Inequalities, Convergence

Expectation an Variance

Covariance and correlation

Covariance mat

Conditional

Characteristic

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence
Convergence typ
Law of large
numbers
Central limit

For discrete variables the generating function is defined as

$$G_X(z) = \sum_i z^{x_i} f_X(x_i).$$

- \rightarrow The characteristic function is $\phi_X(t) = G_X(z = e^{it})$ and the moment generating function is $\psi_X(t) = G_X(z = e^t)$.
- For independent random variables X_1, X_2, \ldots, X_n , the characteristic function, the moment generating function (and for discrete variables, the generating function) for their sum $X = \sum_{i=1}^{n} X_i$ is simply

$$\phi_X(t) = \prod_{i=1}^n \phi_{X_i}(t), \qquad \psi_X(t) = \prod_{i=1}^n \psi_{X_i}(t), \qquad G_X(z) = \prod_{i=1}^n G_{X_i}(z).$$

Expectation, Inequalities, Convergence

xpectation and ariance

Covariance a

correlation

O III

expectation

Characteriet

function

Inequalities

Markov's inequalit

Chebyshev's

Hooffding's inagual

Cauchy-Schwarz

Jensen's inequality

Convergence

Convergence types

numbers

Central IIIII

INEQUALITIES

Markov's inequality

Expectation, Inequalities, Convergence

Markov's inequality

Markov's inequality

Assuming that *X* is a non-negative random variable for which $\mathbb{E}(X)$ exists, for any z < 0 the following inequality holds:

$$P(X>z)\leq \frac{\mathbb{E}(X)}{z}.$$

Chebyshev's inequality

Expectation, Inequalities, Convergence

Expectation an Variance

Covariance and correlation

Covariance mate

Conditional expectation

Characteristic function

Inequalities

Chebyshev's

inequality

Cauchy-Schwarz inequality

Convergence types
Law of large

Central limit

Chebyshev's inequality

Assuming $\mathbb{E}(X) = \mu$ and $\mathbb{V}(X) = \sigma^2$ for a random variable X, the following inequalities hold:

$$P(|X-\mu|\geq z) \leq \frac{\sigma^2}{z^2}$$

$$P\left(\frac{|X-\mu|}{\sigma} \ge z\right) \le \frac{1}{z^2}$$

Ā

Convergence types
Law of large

numbers Central limit

Hoeffding's inequality for Bernoulli random variables

Assume IID Bernoulli distributed random variables
 X₁, X₂,..., X_n ~ Bernoulli(p):

$$P(X_i=1)=p, \qquad P(X_i=0)=1-p, \qquad \forall i.$$

The sample mean is defined as

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i,$$

and **Hoeffding's inequality** states that this is getting "exponentially" close to p if n is increased, since for any $\epsilon > 0$

$$P(|\overline{X_n} - p| \ge \epsilon) \le 2e^{-2n\epsilon^2}$$
.

Expectation, Inequalities, Convergence

Expectation an Variance

Covariance and correlation

Covariance matri

Conditional expectation

Characteristic

Inequalities

Chebyshev's inequality

Hoeffding's inequality

Cauchy-Schwarz inequality Jensen's inequality

Convergence
Convergence types

Central limit

Hoeffding's inequality (in general)

- Assume X₁, X₂,..., X_n are independent random variables, bounded by the intervals [a_i, b_i], (thus, a_i ≤ X_i ≤ b_i).
- · The sample mean is defined as usual,

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i,$$

and according to **Hoeffding's inequality** for any $\epsilon > 0$

$$P(|\overline{X_n} - \mathbb{E}(\overline{X_n})| \ge \epsilon) \le 2 \exp\left(-\frac{2n^2\epsilon^2}{\sum\limits_{i=1}^n (b_i - a_i)^2}\right).$$

Ā

Expectation, Inequalities, Convergence

Variance
Covariance and correlation
Covariance matrix

Inequalities

Markov's inequality Chebyshev's inequality

Hoeffding's inequality
Cauchy-Schwarz
inequality
Jensen's inequality

Convergence type
Law of large
numbers

Proof:

• First we need Hoeffding's Lemma, stating that assuming a random variable X bounded by the interval [a,b] with $\mathbb{E}(X)=0$, for any λ

$$\mathbb{E}(e^{\lambda X}) \le \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right).$$

- Proof of the lemma:
 - Since e^{sx} is a convex function,

$$\forall x \in [a,b]: \quad e^{sx} \le \frac{b-x}{b-a}e^{sa} + \frac{x-a}{b-a}e^{sb}.$$

- Applying $\ensuremath{\mathbb{E}}$ to both sides we obtain

$$\mathbb{E}(e^{sX}) \leq \frac{b - \mathbb{E}(X)}{b - a} e^{sa} + \frac{\mathbb{E}(X) - a}{b - a} e^{sb} = \frac{b}{\mathbb{E}(X) = 0} \frac{b}{b - a} e^{sa} - \frac{a}{b - a} e^{sb} = \left(-\frac{a}{b - a}\right) e^{sa} \left[-\frac{b}{a} + e^{sb - sa}\right] = \left(-\frac{a}{b - a}\right) e^{sa} \left[-\frac{b - a + a}{a} + e^{s(b - a)}\right] =$$

7

Expectation, Inequalities, Convergence

Variance

correlation
Covariance ma

expectation
Characteristic

Inequalities

Markov's inequalit Chebyshev's inequality

Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence
Convergence types
Law of large
numbers

$$\mathbb{E}(e^{sX}) \leq \left(-\frac{a}{b-a}\right)e^{sa}\left[-\frac{b-a+a}{a}+e^{s(b-a)}\right] = \left(-\frac{a}{b-a}\right)e^{sa}\left[-\frac{b-a}{a}-1+e^{s(b-a)}\right].$$

- Let us define $\theta = -\frac{a}{b-a} > 0$, yielding

$$\mathbb{E}(e^{sX}) \leq \left(1 - \theta + \theta e^{s(b-a)}\right) e^{-s\theta(b-a)}.$$

- We further define u = s(b-a) and the function $\varphi : \mathbb{R} \to \mathbb{R}$ with $\varphi(u) = -\theta u + \ln(1 - \theta + \theta e^u)$. This way

$$\mathbb{E}(e^{sX}) \leq e^{\phi(u)}.$$

 We are going to use Taylor's theorem (the Extended Mean Value Theorem), stating that for every u there exists a v between 0 and u such that

$$\varphi(u) = \varphi(0) + u\varphi'(0) + \frac{u^2}{2}\varphi''(v).$$

Expectation, Inequalities, Convergence

Hoeffding's inequality

- The derivatives of $\varphi(u) = -\theta u + \ln(1 - \theta + \theta e^u)$ are

$$\varphi(0) = 0$$

$$\varphi'(0) = -\theta + \frac{\theta e^{u}}{1 - \theta + \theta e^{u}}\Big|_{u=0} = 0$$

$$\varphi''(v) = \underbrace{\frac{\theta e^{v}}{1 - \theta + \theta e^{v}}}_{t} \left(1 - \frac{\theta e^{v}}{1 - \theta + \theta e^{v}}\right) = t(1 - t) \le \frac{1}{4}.$$

Thus, according to Taylor's theorem

$$\varphi(u) \le 0 + 0 \cdot u + \frac{u^2}{2} \frac{1}{4} = \frac{s^2(b-a)^2}{8},$$

proving the lemma

$$\mathbb{E}(e^{sX}) \le e^{\varphi(u)} \le \exp\left(\frac{s^2(b-a)^2}{8}\right).$$

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation

Covariance matr

Conditional expectation

Characteristic function

Inequalities

inequality

Hoeffding's inequality

Cauchy-Schwarz

Cauchy-Schwarz inequality Jensen's inequality

Convergence type: Law of large numbers Central limit Back to the main proof:
 We first point out that equivalently to

$$P(|\overline{X_n} - \mathbb{E}(\overline{X_n})| \ge \epsilon) \le 2 \exp\left(-\frac{2n^2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

also

$$P\left(\overline{X_n} - \mathbb{E}(\overline{X_n}) \ge \epsilon\right) \le \exp\left(-\frac{2n^2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

Taking the expression on the left, for any s > 0 we can write

$$P\left(\overline{X_n} - \mathbb{E}(\overline{X_n}) \ge \epsilon\right) = P\left(e^{s(\overline{X_n} - \mathbb{E}(\overline{X_n}))} > e^{s\epsilon}\right).$$

Since the variable on the left is non-negative, we can use Markov's inequality as

$$P\left(\overline{X_n} - \mathbb{E}(\overline{X_n}) \ge \epsilon\right) = P\left(e^{s\left(\overline{X_n} - \mathbb{E}(\overline{X_n})\right)} > e^{s\epsilon}\right) \le e^{-s\epsilon} \mathbb{E}\left[e^{s\left(\overline{X_n} - \mathbb{E}(\overline{X_n})\right)}\right] = e^{-s\epsilon} \prod_{i=1}^n \mathbb{E}\left[e^{\frac{s}{n}(X_i - \mathbb{E}(X_i))}\right]$$

7

Expectation, Inequalities, Convergence

Expectation an Variance

Covariance and correlation

Covariance matr

Conditional expectation

Characteristic function

Inequalitie

Markov's inequali Chebyshev's inequality

Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence typ Convergence typ Law of large numbers Central limit According to the above

$$P\left(\overline{X_n} - \mathbb{E}(\overline{X_n}) \ge \epsilon\right) \le e^{-s\epsilon} \prod_{i=1}^n \mathbb{E}\left[e^{\frac{s}{n}(X_i - \mathbb{E}(X_i))}\right].$$

We can use Hoeffding's lemma to the term on the right as

$$\begin{split} P\left(\overline{X_n} - \mathbb{E}(\overline{X_n}) \geq \epsilon\right) & \leq e^{-s\epsilon} \prod_{i=1}^n \mathbb{E}\left[e^{\frac{s}{n}(X_i - \mathbb{E}(X_i))}\right] \leq \\ & e^{-s\epsilon} \prod_{i=1}^n e^{\frac{s^2(b_i - a_i)^2}{8n^2}} = \exp\left(-s\epsilon + \frac{s^2}{8n^2} \sum_{i=1}^n (b_i - a_i)^2\right). \end{split}$$

• To obtain the best possible upper bound we have to find the minimum of $g(s) = -s\epsilon + \frac{s^2}{8n^2} \sum_{i=1}^n (b_i - a_i)^2$ as a function of s, which happens to be at $s = 4\epsilon n^2 \left[\sum_{i=1}^n (b_i - a_i)^2\right]^{-1}$. By substituting back we obtain that

$$P\left(\overline{X_n} - \mathbb{E}(\overline{X_n}) \ge \epsilon\right) \le \exp\left(-\frac{2\epsilon^2 n^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

<u>_</u>

Cauchy-Schwarz inequality

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation

Covariance matrix

Conditional expectation

Characteristic function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality

Cauchy–Schwarz inequality Jensen's inequalit

Convergence
Convergence types
Law of large

Central limit

Cauchy-Schwarz inequality

For any random variables X and Y for which $\mathbb{E}(X)$ exists, the following inequality holds:

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}.$$

Proof:

• For any real number a we know that $(Y - aX)^2 \ge 0$, thus,

$$0 \le \mathbb{E}[(Y - aX)^2] = \mathbb{E}[Y^2 - 2aYX + a^2X^2] = \mathbb{E}(Y^2) - 2a\mathbb{E}(YX) + a^2\mathbb{E}(X^2)$$

• Let us choose $a = \frac{\mathbb{E}(XY)}{\mathbb{E}(Y^2)}$:

 $0 \le \mathbb{E}(Y^2) - \frac{\mathbb{E}(XY)^2}{\mathbb{E}(X^2)} \rightarrow 0 \le \mathbb{E}(Y^2)\mathbb{E}(X^2) - \mathbb{E}(XY)^2$

Cauchy-Schwarz inequality

Expectation, Inequalities, Convergence

Expectation and Variance

Covariance and correlation

Covariance matrix

Conditional expectation

Characteristic

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality

Cauchy–Schwarz inequality Jensen's inequality

Convergence
Convergence types
Law of large
numbers

Cauchy-Schwarz inequality

For any random variables X and Y for which $\mathbb{E}(X)$ exists, the following inequality holds:

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}.$$

Proof:

• For any real number a we know that $(Y - aX)^2 \ge 0$, thus,

$$0 \le \mathbb{E}\left[\left(Y - aX\right)^2\right] = \mathbb{E}\left[Y^2 - 2aYX + a^2X^2\right] = \mathbb{E}(Y^2) - 2a\mathbb{E}(YX) + a^2\mathbb{E}(X^2).$$

• Let us choose $a = \frac{\mathbb{E}(XY)}{\mathbb{E}(X^2)}$

$$0 \le \mathbb{E}(Y^2) - \frac{\mathbb{E}(XY)^2}{\mathbb{E}(X^2)} \to 0 \le \mathbb{E}(Y^2)\mathbb{E}(X^2) - \mathbb{E}(XY)^2$$

Cauchy-Schwarz inequality

Expectation, Inequalities, Convergence

Expectation an Variance Covariance and

Covariance and correlation

Covariance matri:

Conditional expectation

Characteristic

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality

Cauchy–Schwarz inequality Jensen's inequality

Convergence
Convergence type
Law of large
numbers

Cauchy-Schwarz inequality

For any random variables X and Y for which $\mathbb{E}(X)$ exists, the following inequality holds:

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}.$$

Proof:

• For any real number a we know that $(Y - aX)^2 \ge 0$, thus,

$$0 \le \mathbb{E}\left[(Y - aX)^2 \right] = \mathbb{E}\left[Y^2 - 2aYX + a^2X^2 \right] = \mathbb{E}(Y^2) - 2a\mathbb{E}(YX) + a^2\mathbb{E}(X^2).$$

• Let us choose $a = \frac{\mathbb{E}(XY)}{\mathbb{E}(X^2)}$:

$$0 \le \mathbb{E}(Y^2) - \frac{\mathbb{E}(XY)^2}{\mathbb{E}(X^2)} \rightarrow 0 \le \mathbb{E}(Y^2)\mathbb{E}(X^2) - \mathbb{E}(XY)^2$$

Jensen's inequality

Expectation, Inequalities, Convergence

Expectation an Variance

Covariance and correlation
Covariance matri
Conditional expectation
Characteristic function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy—Schwarz inequality

Jensen's inequality
Convergence
Convergence types
Law of large

Jensen's inequality

If f(x) is a convex function, and g(x) is a concave function then

$$\mathbb{E}[f(X)] \ge f(\mathbb{E}[x]),$$

$$\mathbb{E}[g(X)] \leq g(\mathbb{E}[x]).$$

Proof

Let h(x) = ax + b be a line tangent to f(x) at the point $x = \mathbb{E}(X)$. Since f(x) is convex, it lies above this line. Thus,

$$\mathbb{E}[f(X)] \ge \mathbb{E}[h(x)] = \mathbb{E}[ax + b] = a\mathbb{E}(X) + b = h(x = \mathbb{E}[X]) = f(\mathbb{E}[X]).$$

 Consequence: E.g., since f(x) = x² and f(x) = 1/x are convex, for any variable in general

$$\mathbb{E}(X^2) \ge (\mathbb{E}[X])^2$$
, $\mathbb{E}(1/X) \ge \frac{1}{\mathbb{E}(X)}$

numbers
Central limit

Jensen's inequality

If f(x) is a convex function, and g(x) is a concave function then

$$\mathbb{E}[f(X)] \ge f(\mathbb{E}[x]),$$

$$\mathbb{E}[g(X)] \leq g(\mathbb{E}[x]).$$

· Proof:

Let h(x) = ax + b be a line tangent to f(x) at the point $x = \mathbb{E}(X)$. Since f(x) is convex, it lies above this line. Thus,

$$\mathbb{E}[f(X)] \geq \mathbb{E}[h(x)] = \mathbb{E}[ax + b] = a\mathbb{E}(X) + b = h(x = \mathbb{E}[X]) = f(\mathbb{E}[X]).$$

• Consequence: E.g., since $f(x) = x^2$ and f(x) = 1/x are convex, for any variable in general

$$\mathbb{E}(X^2) \ge (\mathbb{E}[X])^2$$
, $\mathbb{E}(1/X) \ge \frac{1}{\mathbb{E}(X)}$

Convergence type
Convergence type
Law of large
numbers
Central limit
theorems

Jensen's inequality

If f(x) is a convex function, and g(x) is a concave function then

$$\mathbb{E}[f(X)] \ge f(\mathbb{E}[x]),$$

$$\mathbb{E}[g(X)] \leq g(\mathbb{E}[x]).$$

· Proof:

Let h(x) = ax + b be a line tangent to f(x) at the point $x = \mathbb{E}(X)$. Since f(x) is convex, it lies above this line. Thus,

$$\mathbb{E}[f(X)] \geq \mathbb{E}[h(x)] = \mathbb{E}[ax + b] = a\mathbb{E}(X) + b = h(x = \mathbb{E}[X]) = f(\mathbb{E}[X]).$$

• Consequence: E.g., since $f(x) = x^2$ and f(x) = 1/x are convex, for any variable in general

$$\mathbb{E}(X^2) \ge (\mathbb{E}[X])^2, \qquad \mathbb{E}(1/X) \ge \frac{1}{\mathbb{E}(X)}$$

Expectation, Inequalities, Convergence

rpectation and riance

Covariance ar

Covariance ma

Conditional

expectation

Characteristi

Inequalities

Markov's inequal

Chebyshev's

inequality

Hoeffding's inequalit

inequality

Jensen's inequali

Convergence

Law of laws

numbers

Central limi

CONVERGENCE

Expectation, Inequalities, Convergence

Expectation an Variance

correlation
Covariance mate
Conditional
expectation
Characteristic
function

Inequalities

Chebyshev's inequality
Hoeffding's inequality
Cauchy–Schwarz inequality

Convergence Convergence types

Law of large numbers

Convergence in probability

If $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of random variables, then X_n is converging to X in probability (or weakly) if for every $\epsilon > 0$

$$\lim_{n\to\infty}P(|X_n-X|>\epsilon)\to 0.$$

This is usually denoted by $X_n \stackrel{P}{\longrightarrow} X$.

Expectation, Inequalities, Convergence

Expectation ar Variance

correlation

Covariance matri

Conditional

expectation

Characteristic

function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence types

Law of large numbers Central limit

Convergence in probability

If $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of random variables, then X_n is converging to X in probability (or weakly) if for every $\epsilon > 0$

$$\lim_{n\to\infty}P\big(\big|X_n-X\big|>\epsilon\big)\to 0.$$

This is usually denoted by $X_n \stackrel{P}{\longrightarrow} X$.

Almost sure convergence

If $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of random variables, then X_n is converging to X almost surely (or almost everywhere, or with probability 1, or strongly) if

$$P\left(\lim_{n\to\infty}X_n=X\right)=1.$$

This is usually denoted by $X_n \stackrel{\text{a.s.}}{\longrightarrow} X$.

Expectation, Inequalities, Convergence

Expectation an Variance

correlation
Covariance mat
Conditional
expectation
Characteristic

Inequalities

Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality

Convergence Convergence types

Law of large numbers

Convergence in distribution

If $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of random variable, where the corresponding CDFs are denoted by F_n , then X_n is converging to X (with CDF given by F(x)) in distribution if

$$\lim_{n\to\infty}F_n(x)=F(x)$$

at all x where F(x) is continuous. This is usually denoted by $X_n \stackrel{d}{\longrightarrow} X$

Ā

Expectation, Inequalities, Convergence

Expectation ar Variance Covariance and correlation

correlation
Covariance matri
Conditional
expectation
Characteristic
function

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence types

Law of large numbers Central limit theorems

Convergence in distribution

If $X_1, X_2, \dots, X_n, \dots$ is a sequence of random variable, where the corresponding CDFs are denoted by F_n , then X_n is converging to X (with CDF given by F(x)) in distribution if

$$\lim_{n\to\infty}F_n(x)=F(x)$$

at all *x* where F(x) is continuous. This is usually denoted by $X_n \stackrel{d}{\longrightarrow} X$

Convergence in mean

If $X_1, X_2, ..., X_n, ...$ is a sequence of random variable and $r \ge 1$, then X_n is converging to X in the r-th mean (or in the L^r -th norm) if

$$\lim_{n\to\infty}\mathbb{E}(|X_n-X|^r)=0.$$

This is usually denoted as $X_n \stackrel{L^r}{\longrightarrow} X$.

Expectation, Inequalities, Convergence

Expectation and Variance

correlation

Covariance ma

Conditional

expectation

Characteris

Inequalities

Chebyshev's

inequality

Hoeffding's inequa Cauchy-Schwarz inequality

Jensen's inequa

Convergence types

Convergence types

numbers

Chain of implications:

$$\begin{array}{cccc} \overset{L^s}{\longrightarrow} & \underset{s>r\geq 1}{\Longrightarrow} & \overset{L^r}{\longrightarrow} & \\ & & & \downarrow & \\ & & \downarrow & \\ & \text{a.s.} & \Rightarrow & \overset{P}{\longrightarrow} & \Rightarrow & \overset{d}{\longrightarrow} & \end{array}$$

Law of large numbers

Expectation, Inequalities, Convergence

Expectation an Variance

correlation
Covariance matri
Conditional
expectation

Inequalities

Chebyshev's inequality

Hoeffding's inequali Cauchy-Schwarz inequality

Convergence type

Law of large

numbers Central limit

Law of large numbers

The weak law of large number states that if X_1, X_2, \dots, X_n are IID where $\mathbb{E}(X_i) = \mu$, then

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i \stackrel{P}{\longrightarrow} \mu.$$

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence typ

Law of large

numbers Central limit

Law of large numbers

The weak law of large number states that if X_1, X_2, \dots, X_n are IID where $\mathbb{E}(X_i) = \mu$, then

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i \stackrel{P}{\longrightarrow} \mu.$$

Proof:

Assuming that $\mathbb{V}(X_i) = \sigma^2 < \infty$ we can use Chebyshev's inequality as

$$P(|\overline{X_n} - \mu| > \epsilon) \le \frac{\mathbb{V}(\overline{X_n})}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \xrightarrow[n \to \infty]{} 0$$

Hoeffding's inequality
Cauchy-Schwarz
inequality
Jensen's inequality

Convergence
Convergence types
Law of large

Central limit

The Central limit theorem

Let $X_1, X_2, \ldots, X_n, \ldots$ be IID random variables with $\mathbb{E}(X_i) = \mu$ and $\mathbb{V}(X_i) = \sigma^2$. Then the transformed variables

$$Z_n = \frac{\sqrt{n}(\overline{X_n} - \mu)}{\sigma} \stackrel{d}{\longrightarrow} Z \sim N(0, 1).$$

In other words,

$$\lim_{n\to\infty}P\big(Z_n\leq z\big)=\Phi\big(z\big)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^z e^{-\frac{x^2}{2}}dx.$$

Central limit theorem for heavy tailed distributions

Expectation, Inequalities, Convergence

Expectation and

Covariance and correlation
Covariance matri:
Conditional expectation

Inequalities

Markov's inequality Chebyshev's inequality Hoeffding's inequality Cauchy-Schwarz inequality Jensen's inequality

Convergence Convergence types Law of large

Central limit

Central limit theorem for heavy tailed distributions

If X_1, X_2, \dots, X_n are IID variables with a PDF such that

$$\rho(x) \approx \frac{C_0}{|x|^{\mu+1}}$$

for |x| >> 1 where $1 \le \mu \le 2$, then the transformed variable

$$Y_n = \frac{1}{n^{1/\mu}} \sum_{i=1}^n (X_i - \mathbb{E}(X_i))$$

will converge in distribution to a Levy distribution, $Y_n \stackrel{d}{\longrightarrow} Y \sim \text{Levy}(\mu)$. which can be most easily specified by its characteristic function.

$$\phi_Y(t) = e^{-C|t|^{\mu}},$$

having also a heavy tail,

$$\rho_Y(y) \approx \frac{C_1}{|y|^{\mu+1}}$$

for y >> 1.

<u>_</u>