

Chebyshev's inequality

From Wikipedia, the free encyclopedia

In probability theory, **Chebyshev's inequality** (also spelled as **Tchebysheff's inequality**, Russian: Неравенство Чебышёва) guarantees that in any probability distribution, "nearly all" values are close to the mean — the precise statement being that no more than $1/k^2$ of the distribution's values can be more than k standard deviations away from the mean (or equivalently, at least $1-1/k^2$ of the distribution's values are within k standard deviations of the mean). The rule is often called Chebyshev's theorem, about the range of standard deviations around the mean, in statistics. The inequality has great utility because it can be applied to completely arbitrary distributions (unknown except for mean and variance). For example, it can be used to prove the weak law of large numbers.

In practical usage, in contrast to the 68-95-99.7% rule, which applies to normal distributions, under Chebyshev's inequality a minimum of just 75% of values must lie within two standard deviations of the mean and 89% within three standard deviations.^{[1][2]}

The term *Chebyshev's inequality* may also refer to Markov's inequality, especially in the context of analysis.

Contents

- 1 History
 - 2 Statement
 - 2.1 Probabilistic statement
 - 2.2 Measure-theoretic statement
 - 3 Example
 - 4 Sharpness of bounds
 - 5 Proof (of the two-sided version)
 - 5.1 Probabilistic proof
 - 5.2 Measure-theoretic proof
 - 6 Extensions
 - 6.1 Asymmetric two-sided case
 - 6.2 Bivariate case
 - 6.3 Two correlated variables
 - 6.4 Multivariate case
 - 6.5 Vector version
 - 6.6 Infinite dimensions
 - 6.7 Higher moments
 - 6.8 Exponential version
 - 6.9 Inequalities for bounded variables
 - 7 Finite samples
 - 7.1 Dependence of sample size
 - 8 Samuelson's inequality
 - 9 Sharpened bounds
 - 9.1 Standardised variables
 - 9.2 Semivariances
 - 9.3 Selberg's inequality
 - 9.4 Cantelli's inequality
 - 9.5 Bhattacharyya's inequality
 - 9.6 Mitzenmacher and Upfal's inequality
 - 10 Related inequalities
 - 10.1 Zelen's inequality

- - 10.2 He, Zhang and Zhang's inequality
 - 10.3 Hoeffding's lemma
 - 10.4 Van Zuijlen's bound
- - 11 Unimodal distributions
 - - 11.1 Unimodal symmetrical distributions
 - 11.2 Notes
 - 12 Bounds for specific distributions
- - 13 Zero means
 - - 13.1 Unit variance
 - 14 Integral Chebyshev inequality
 - - 14.1 Other inequalities
- - 15 Haldane's transformation
- - 16 Notes
- - 17 See also
- - 18 References
- - 19 Further reading
- - 20 External links

History

The theorem is named after Russian mathematician Pafnuty Chebyshev, although it was first formulated by his friend and colleague Irénée-Jules Bienaymé.^{[3]:98} The theorem was first stated without proof by Bienaymé in 1853^[4] and later proved by Chebyshev in 1867.^[5] His student Andrey Markov provided another proof in his 1884 Ph.D. thesis.^[6]

Statement

Chebyshev's inequality is usually stated for random variables, but can be generalized to a statement about measure spaces.

Probabilistic statement

Let X (integrable) be a random variable with finite expected value μ and finite non-zero variance σ^2 . Then for any real number $k > 0$,

$$\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Only the case $k > 1$ is useful. When $k \leq 1$ the right hand : $\frac{1}{k^2} \geq 1$, and the inequality is trivial as all probabilities are ≤ 1 .

As an example, using $k = \sqrt{2}$ shows that the probability that values lie outside the interval $(\mu - \sqrt{2}\sigma, \mu + \sqrt{2}\sigma)$ does not exceed 1/2.

Because it can be applied to completely arbitrary distributions (unknown except for mean and variance), the inequality generally gives a poor bound compared to what might be deduced if more aspects are known about the distribution involved.

k	Min % within k standard deviations of mean	Max % beyond k standard deviations from mean
1	0%	100%
$\sqrt{2}$	50%	50%
1.5	55.56%	44.44%
2	75%	25%
3	88.8889%	11.1111%
4	93.75%	6.25%
5	96%	4%
6	97.2222%	2.7778%
7	97.9592%	2.0408%
8	98.4375%	1.5625%
9	98.7654%	1.2346%
10	99%	1%

Measure-theoretic statement

Let (X, Σ, μ) be a measure space, and let f be an extended real-valued measurable function defined on X . Then for any real number $t > 0$ and $0 < p < \infty$,^[7]

$$\mu(\{x \in X : |f(x)| \geq t\}) \leq \frac{1}{t^p} \int_{|f| \geq t} |f|^p \, d\mu.$$

More generally, if g is an extended real-valued measurable function, nonnegative and nondecreasing on the range of f , then

$$\mu(\{x \in X : f(x) \geq t\}) \leq \frac{1}{g(t)} \int_X g \circ f \, d\mu.$$

The previous statement then follows by defining $g(t)$ as t^p if $t \geq 0$ and 0 otherwise, and taking $|f|$ instead of f .

Example

Suppose we randomly select a journal article from a source with an average of 1000 words per article, with a standard deviation of 200 words. We can then infer that the probability that it has between 600 and 1400 words (i.e. within $k = 2$ standard deviations of the mean) must be at least 75%, because there is no more than $1/k^2 = 1/4$ chance to be outside that range, by Chebyshev's inequality. But if we additionally know that the distribution is normal, we can say that is a 75% chance the word count is between 770 and 1230 (which is an even tighter bound).

Sharpness of bounds

As shown in the example above, the theorem typically provides rather loose bounds. However, these bounds cannot in general (remaining true for arbitrary distributions) be improved upon. The bounds are sharp for the following example: for any $k \geq 1$,

$$X = \begin{cases} -1, & \text{with probability } \frac{1}{2k^2} \\ 0, & \text{with probability } 1 - \frac{1}{k^2} \\ 1, & \text{with probability } \frac{1}{2k^2} \end{cases}$$

For this distribution, mean $\mu = 0$ and standard deviation $\sigma = \frac{1}{k}$, so

$$\Pr(|X - \mu| \geq k\sigma) = \Pr(|X| \geq 1) = \frac{1}{k^2}.$$

Chebyshev's inequality is an equality for precisely those distributions that are a linear transformation of this example.

Proof (of the two-sided version)

Probabilistic proof

Markov's inequality states that for any real-valued random variable Y and any positive number a , we have $\Pr(|Y| > a) \leq E(|Y|)/a$. One way to prove Chebyshev's inequality is to apply Markov's inequality to the random variable $Y = (X - \mu)^2$ with $a = (k\sigma)^2$.

It can also be proved directly. For any event A , let I_A be the indicator random variable of A , i.e. I_A equals 1 if A occurs and 0 otherwise. Then

$$\begin{aligned}
 \Pr(|X - \mu| \geq k\sigma) &= \mathbb{E}(I_{|X - \mu| \geq k\sigma}) \\
 &= \mathbb{E}\left(I_{\left(\frac{X - \mu}{k\sigma}\right)^2 \geq 1}\right) \\
 &\leq \mathbb{E}\left(\left(\frac{X - \mu}{k\sigma}\right)^2\right) \\
 &= \frac{1}{k^2} \frac{\mathbb{E}((X - \mu)^2)}{\sigma^2} \\
 &= \frac{1}{k^2}.
 \end{aligned}$$

The direct proof shows why the bounds are quite loose in typical cases: the number 1 to the right of " \geq " is replaced by $[(X - \mu)/(k\sigma)]^2$ to the left of " \geq " whenever the latter exceeds 1. In some cases it exceeds 1 by a very wide margin.

Measure-theoretic proof

Fix t and let A_t be defined as $A_t = \{x \in X \mid f(x) \geq t\}$, and let 1_{A_t} be the indicator function of the set A_t . Then, it is easy to check that, for any x ,

$$0 \leq g(t)1_{A_t} \leq g(f(x))1_{A_t},$$

since g is nondecreasing on the range of f , and therefore,

$$\begin{aligned}
 g(t)\mu(A_t) &= \int_X g(t)1_{A_t} d\mu \\
 &\leq \int_{A_t} g \circ f d\mu \\
 &\leq \int_X g \circ f d\mu.
 \end{aligned}$$

The desired inequality follows from dividing the above inequality by $g(t)$.

Extensions

Several extensions of Chebyshev's inequality have been developed.

Asymmetric two-sided case

An asymmetric two-sided version of this inequality is also known.^[8]

When the distribution is known to be symmetric for any $k_1 + k_2 = 0$

$$\Pr(k_1 < X < k_2) \geq 1 - \frac{4\sigma^2}{(k_2 - k_1)^2}$$

where σ^2 is the variance.

Similarly when the distribution is asymmetric or is unknown and $k_1 + k_2 = 2\mu$

$$\Pr(k_1 < X < k_2) \geq \frac{4[(\mu - k_1)(k_2 - \mu) - \sigma^2]}{(k_2 - k_1)^2},$$

where σ^2 is the variance and μ is the mean.

Bivariate case

A version for the bivariate case is known.^[9]

Let X_1, X_2 be two random variables with means μ_1, μ_2 and finite variances σ_1, σ_2 respectively. Then

$$\Pr(k_{11} \leq X_1 \leq k_{12}, k_{21} \leq X_2 \leq k_{22}) \geq 1 - \sum T_i$$

where for $i = 1, 2$,

$$T_i = \frac{4\sigma_i^2 + [2\mu_i - (k_{i1} + k_{i2})]^2}{(k_{i2} - k_{i1})^2}.$$

Two correlated variables

Berge derived an inequality for two correlated variables X_1, X_2 .^[10] Let ρ be the correlation coefficient between X_1 and X_2 and let σ_i^2 be the variance of X_i . Then

$$\Pr \left(\bigcap_{i=1}^2 \left[\frac{|X_i - \mu_i|}{\sigma_i} < k \right] \right) \geq 1 - \frac{1 + \sqrt{1 - \rho^2}}{k^2}.$$

Lal later obtained an alternative bound^[11]

$$\Pr \left(\bigcap_{i=1}^2 \left[\frac{|X_i - \mu_i|}{\sigma_i} \leq k_i \right] \right) \geq 1 - \frac{k_1^2 + k_2^2 + \sqrt{(k_1^2 + k_2^2)^2 - 4k_1^2 k_2^2 \rho}}{2(k_1 k_2)^2}$$

Isii derived a further generalisation.^[12] Let

$$Z = \Pr ((-k_1 < X_1 < k_2) \cap (-k_1 < X_2 < k_2)), \quad 0 < k_1 \leq k_2.$$

and define:

$$\lambda = \frac{k_1(1 + \rho) + \sqrt{(1 - \rho^2)(k_1^2 + \rho)}}{2k_1 - 1 + \rho}$$

There are now three cases.

- **Case A:** If $2k_1^2 > 1 - \rho$ and $k_2 - k_1 \geq 2\lambda$ then

$$Z \leq \frac{2\lambda^2}{2\lambda^2 + 1 + \rho}.$$

- **Case B:** If the conditions in case A are not met but $k_1 k_2 \geq 1$ and

$$2(k_1 k_2 - 1)^2 \geq 2(1 - \rho^2) + (1 - \rho)(k_2 - k_1)^2$$

then

$$Z \leq \frac{(k_2 - k_1)^2 + 4 + \sqrt{16(1 - \rho^2) + 8(1 - \rho)(k_2 - k_1)}}{(k_1 + k_2)^2}.$$

- **Case C:** If none of the conditions in cases A or B are satisfied then there is no universal bound other than 1.

Multivariate case

The general case is known as the Birnbaum–Raymond–Zuckerman inequality after the authors who proved it for two dimensions.^[13]

$$\Pr \left(\sum_{i=1}^n \frac{(X_i - \mu_i)^2}{\sigma_i^2 t_i^2} \geq k^2 \right) \leq \frac{1}{k^2} \sum_{i=1}^n \frac{1}{t_i^2}$$

where X_i is the i -th random variable, μ_i is the i -th mean and σ_i^2 is the i -th variance.

If the variables are independent this inequality can be sharpened.^[14]

$$\Pr \left(\bigcap_{i=1}^n \frac{|X_i - \mu_i|}{\sigma_i} \leq k_i \right) \geq \prod \left(1 - \frac{1}{k_i^2} \right)$$

Olkin and Pratt derived an inequality for n correlated variables.^[15]

$$\Pr \left(\bigcap_{i=1}^n \frac{|X_i - \mu_i|}{\sigma_i} < k_i \right) \geq 1 - \frac{1}{n^2} \left(\sqrt{U} + \sqrt{n-1} \sqrt{n \sum_i \frac{1}{k_i^2} - u} \right)^2$$

where the sum is taken over the n variables and

$$u = \sum_{i=1}^n \frac{1}{k_i^2} + 2 \sum_{i=1}^n \sum_{j < i} \frac{\rho_{ij}}{k_i k_j}$$

where ρ_{ij} is the correlation between X_i and X_j .

Olkin and Pratt's inequality was subsequently generalised by Godwin.^[16]

Vector version

Ferentinos^[9] has shown that for a vector $X = (x_1, x_2, \dots)$ with mean $\mu = (\mu_1, \mu_2, \dots)$, variance $\sigma^2 = (\sigma_1^2, \sigma_2^2, \dots)$ and the Euclidean norm $\| \cdot \|$ that

$$\Pr(\|X - \mu\| \geq k\|\sigma\|) \leq \frac{1}{k^2}.$$

A second related inequality has also been derived by Chen.^[17] Let N be the dimension of the stochastic vector X and let $E(X)$ be the mean of X . Let S be the covariance matrix and $k > 0$. Then

$$\Pr((X - E(X))^T S^{-1} (X - E(X)) < k) \geq 1 - \frac{N}{k}$$

where Y^T is the transpose of Y .

Infinite dimensions

There is a straightforward extension of the vector version of Chebyshev's inequality to infinite dimensional settings. Let X be a random variable which takes values in a Fréchet space \mathcal{X} (equipped with seminorms $\|\cdot\|_\alpha$). This includes most common settings of vector-valued random variables, e.g., when \mathcal{X} is a Banach space (equipped with a single norm), a Hilbert space, or the finite-dimensional setting as described above.

Suppose that X is of "strong order two", meaning that

$$E(\|X\|_\alpha^2) < \infty$$

for every seminorm $\|\cdot\|_\alpha$. This is a generalization of the requirement that X have finite variance, and is necessary for this strong form of Chebyshev's inequality in infinite dimensions. The terminology "strong order two" is due to Vakhania.^[18]

Let $\mu \in \mathcal{X}$ be the Pettis integral of X (i.e., the vector generalization of the mean), and let

$$\sigma_\alpha := \sqrt{E\|X - \mu\|_\alpha^2}$$

be the standard deviation with respect to the seminorm $\|\cdot\|_\alpha$. In this setting we can state the following:

$$\textbf{General Version of Chebyshev's Inequality.} \forall k > 0 : \quad \Pr(\|X - \mu\|_\alpha \geq k\sigma_\alpha) \leq \frac{1}{k^2}.$$

Proof. The proof is straightforward, and essentially the same as the finitary version. If $\sigma_\alpha = 0$, then X is constant (and equal to μ) almost surely, so the inequality is trivial.

If

$$\|X - \mu\|_\alpha \geq k\sigma_\alpha^2$$

then $\|X - \mu\|_\alpha > 0$, so we may safely divide by $\|X - \mu\|_\alpha$. The crucial trick in Chebyshev's inequality is to recognize that $1 = \frac{\|X - \mu\|_\alpha^2}{\|X - \mu\|_\alpha^2}$.

The following calculations complete the proof:

$$\begin{aligned} \Pr(\|X - \mu\|_\alpha \geq k\sigma_\alpha) &= \int_{\Omega} \mathbf{1}_{\|X - \mu\|_\alpha \geq k\sigma_\alpha} d\Pr \\ &= \int_{\Omega} \left(\frac{\|X - \mu\|_\alpha^2}{\|X - \mu\|_\alpha^2} \right) \cdot \mathbf{1}_{\|X - \mu\|_\alpha \geq k\sigma_\alpha} d\Pr \\ &\leq \int_{\Omega} \left(\frac{\|X - \mu\|_\alpha^2}{(k\sigma_\alpha)^2} \right) \cdot \mathbf{1}_{\|X - \mu\|_\alpha \geq k\sigma_\alpha} d\Pr \\ &\leq \frac{1}{k^2\sigma_\alpha^2} \int_{\Omega} \|X - \mu\|_\alpha^2 d\Pr & \mathbf{1}_{\|X - \mu\|_\alpha \geq k\sigma_\alpha} &\leq 1 \\ &= \frac{1}{k^2\sigma_\alpha^2} (E\|X - \mu\|_\alpha^2) \\ &= \frac{1}{k^2\sigma_\alpha^2} (\sigma_\alpha^2) \\ &= \frac{1}{k^2} \end{aligned}$$

Higher moments

An extension to higher moments is also possible:

$$\Pr(|X - E(X)| \geq k E(|X - E(X)|^n)^{\frac{1}{n}}) \leq \frac{1}{k^n}, \quad k > 0, n \geq 2.$$

Exponential version

A related inequality sometimes known as the exponential Chebyshev's inequality^[19] is the inequality

$$\Pr(X \geq \varepsilon) \leq e^{-t\varepsilon} \mathbb{E}(e^{tX}), \quad t > 0.$$

Let $K(x, t)$ be the cumulant generating function,

$$K(x, t) = \log(\mathbb{E}(e^{tx})).$$

Taking the Legendre–Fenchel transformation of $K(x, t)$ and using the exponential Chebyshev's inequality we have

$$-\log(\Pr(X \geq \varepsilon)) \leq \sup_t (t\varepsilon - K(x, t)).$$

This inequality may be used to obtain exponential inequalities for unbounded variables.^[20]

Inequalities for bounded variables

If $P(x)$ has finite support based on the interval $[a, b]$, let $M = \max(|a|, |b|)$ where $|x|$ is the absolute value of x . If the mean of $P(x)$ is zero then for all $k > 0$ ^[21]

$$\frac{\mathbb{E}(|X|^r) - k^r}{M^r} \leq \Pr(|X| \geq k) \leq \frac{\mathbb{E}(|X|^r)}{k^r}.$$

The second of these inequalities with $r = 2$ is the Chebyshev bound. The first provides a lower bound for the value of $P(x)$.

Sharp bounds for a bounded variate have been derived by Niemitalo^[22]

Let $0 \leq X \leq M$ where $M > 0$. Then

■ Case 1:

$$\Pr(X < k) = 0 \quad \text{if} \quad \mathbb{E}(X) > k \quad \text{and} \quad \mathbb{E}(X^2) < k \mathbb{E}(X) + M \mathbb{E}(X) - kM$$

■ Case 2:

$$\Pr(X < k) \geq 1 - \frac{k \mathbb{E}(X) + M \mathbb{E}(X) - \mathbb{E}(X^2)}{kM} \quad \text{if} \quad \begin{cases} \mathbb{E}(X) > k & \text{and} & \mathbb{E}(X^2) \geq k \mathbb{E}(X) + M \mathbb{E}(X) - kM \\ \text{or} \\ \mathbb{E}(X) \leq k & \text{and} & \mathbb{E}(X^2) \geq k \mathbb{E}(X) \end{cases}$$

■ Case 3:

$$\Pr(X < k) \geq \frac{\mathbb{E}(X)^2 - 2k \mathbb{E}(X) + k^2}{\mathbb{E}(X^2) - 2k \mathbb{E}(X) + k^2} \quad \text{if} \quad \mathbb{E}(X) \leq k \quad \text{and} \quad \mathbb{E}(X^2) < k \mathbb{E}(X)$$

Finite samples

Saw *et al* extended Chebyshev's inequality to cases where the population mean and variance are not known and may not exist, but you want to use the sample mean and sample standard deviation from N samples to bound the expected value of a new drawing from the same distribution.^[23]

$$P(|X - m| \geq ks) \leq \frac{g_{N+1} \left(\frac{Nk^2}{N-1+k^2} \right)}{N+1} \left(\frac{N}{N+1} \right)^{1/2}$$

where X is a random variable which we have sampled N times, m is the sample mean, k is a constant and s is the sample standard deviation. $g(x)$ is defined as follows:

Let $x \geq 1$, $Q = N + 1$, and R be the greatest integer less than Q/x . Let

$$a^2 = \frac{Q(Q-R)}{1+R(Q-R)}.$$

Now

$$\begin{aligned} g_Q(x) &= R && \text{if } R \text{ is even} \\ g_Q(x) &= R && \text{if } R \text{ is odd and } x < a^2 \\ g_Q(x) &= R - 1 && \text{if } R \text{ is odd and } x \geq a^2. \end{aligned}$$

This inequality holds even when the population moments do not exist, and when the sample is only weakly exchangeably distributed; this criterion is met for randomised sampling. A table of values for the Saw–Yang–Mo inequality for finite sample sizes ($n < 100$) has been determined by Konijn.^[24] The table allows the calculation of various confidence intervals for the mean, based on multiples, C , of the standard error of the mean as calculated from the sample. For example, Konijn shows that for $n = 59$, the 95 per cent confidence interval for the mean m is $(m - Cs, m + Cs)$ where $C = 4.447 \times 1.006 = 4.47$ (this is 2.28 times larger than the value found on the assumption of Normality showing the loss on precision resulting from ignorance of the precise nature of the distribution).

Kabán gives a somewhat less complex version of this inequality.^[25]

$$P(|X - m| \geq ks) \leq \frac{1}{[N(N+1)]^{1/2}} \left[\left(\frac{N-1}{k^2} + 1 \right) \right]$$

If the standard deviation is a multiple of the mean then a further inequality can be derived,^[25]

$$P(|X - m| \geq ks) \leq \frac{N-1}{N} \frac{1}{k^2} \frac{s^2}{m^2} + \frac{1}{N}.$$

A table of values for the Saw–Yang–Mo inequality for finite sample sizes ($n < 100$) has been determined by Konijn.^[24]

For fixed N and large m the Saw–Yang–Mo inequality is approximately^[26]

$$P(|X - m| \geq ks) \leq \frac{1}{N+1}.$$

Beasley *et al* have suggested a modification of this inequality^[26]

$$P(|X - m| \geq ks) \leq \frac{1}{k^2(N+1)}.$$

In empirical testing this modification is conservative but appears to have low statistical power. Its theoretical basis currently remains unexplored.

Dependence of sample size

The bounds these inequalities give on a finite sample are less tight than those the Chebyshev inequality gives for a distribution. To illustrate this let the sample size $n = 100$ and let $k = 3$. Chebyshev's inequality states that at most approximately 11.11% of the distribution will lie outside these limits. Kabán's version of the inequality for a finite sample states that at most approximately 12.05% of the sample lies outside these limits. The dependence of the confidence intervals on sample size is further illustrated below.

For $N = 10$, the 95% confidence interval is approximately ± 13.5789 standard deviations.

For $N = 100$ the 95% confidence interval is approximately ± 4.9595 standard deviations; the 99% confidence interval is approximately ± 140.0 standard deviations.

For $N = 500$ the 95% confidence interval is approximately ± 4.5574 standard deviations; the 99% confidence interval is approximately ± 11.1620 standard deviations.

For $N = 1000$ the 95% and 99% confidence intervals are approximately ± 4.5141 and approximately ± 10.5330 standard deviations respectively.

The Chebyshev inequality for the distribution gives 95% and 99% confidence intervals of approximately ± 4.472 standard deviations and ± 10 standard deviations respectively.

Samuelson's inequality

Although Chebyshev's inequality is the best possible bound for an arbitrary distribution, this is not necessarily true for finite samples. Samuelson's inequality states that all values of a sample will lie within $\sqrt{N-1}$ standard deviations of the mean. Chebyshev's bound improves as the sample size increases.

When $N = 10$, Samuelson's inequality states that all members of the sample lie within 3 standard deviations of the mean: in contrast Chebyshev's states that 99.5% of the sample lies within 13.5789 standard deviations of the mean.

When $N = 100$, Samuelson's inequality states that all members of the sample lie within approximately 9.9499 standard deviations of the mean: Chebyshev's states that 99% of the sample lies within 10 standard deviations of the mean.

When $N = 500$, Samuelson's inequality states that all members of the sample lie within approximately 22.3383 standard deviations of the mean: Chebyshev's states that 99% of the sample lies within 10 standard deviations of the mean.

Sharpened bounds

Chebyshev's inequality is important because of its applicability to any distribution. As a result of its generality it may not (and usually does not) provide as sharp a bound as alternative methods that can be used if the distribution of the random variable is known. To improve the sharpness of the bounds provided by Chebyshev's inequality a number of methods have been developed; for a review see eg.^[27]

Standardised variables

Sharpened bounds can be derived by first standardising the random variable.^[28]

Let X be a random variable with finite variance $\text{Var}(x)$. Let Z be the standardised form defined as

$$Z = \frac{X - \text{E}(X)}{\text{Var}(X)^{1/2}}.$$

Cantelli's lemma is then

$$P(Z \geq k) \leq \frac{1}{1 + k^2}.$$

This inequality is sharp and is attained by k and $-1/k$ with probability $1/(1+k^2)$ and $k^2/(1+k^2)$ respectively.

If $k > 1$ and the distribution of X is symmetric then we have

$$P(Z \geq k) \leq \frac{1}{2k^2}.$$

Equality holds if and only if $Z = -k$, 0 or k with probabilities $1/2k^2$, $1 - 1/k^2$ and $1/2k^2$ respectively.^[28] An extension to a two-sided inequality is also possible.

Let $u, v > 0$. Then we have^[28]

$$P(Z \leq -u \text{ or } Z \geq v) \leq \frac{4 + (u - v)^2}{(u + v)^2}.$$

Semivariances

An alternative method of obtaining sharper bounds is through the use of semivariances (partial moments). The upper (σ_+^2) and lower (σ_-^2) semivariances are defined

$$\sigma_+^2 = \frac{\sum (x - m)^2}{n - 1}$$

$$\sigma_-^2 = \frac{\sum (m - x)^2}{n - 1}$$

where m is the arithmetic mean of the sample, n is the number of elements in the sample and the sum for the upper (lower) semivariance is taken over the elements greater (less) than the mean.

The variance of the sample is the sum of the two semivariances

$$\sigma^2 = \sigma_+^2 + \sigma_-^2.$$

In terms of the lower semivariance Chebyshev's inequality can be written^[29]

$$\Pr(x \leq m - a\sigma_-) \leq \frac{1}{a^2}.$$

Putting

$$a = \frac{k\sigma}{\sigma_-}.$$

Chebyshev's inequality can now be written

$$\Pr(x \leq m - k\sigma) \leq \frac{1}{k^2} \frac{\sigma_-^2}{\sigma^2}.$$

A similar result can also be derived for the upper semivariance.

If we put

$$\sigma_u^2 = \max(\sigma_-^2, \sigma_+^2),$$

Chebyshev's inequality can be written

$$\Pr(|x \leq m - k\sigma|) \leq \frac{1}{k^2} \frac{\sigma_u^2}{\sigma^2}.$$

Because $\sigma_u^2 \leq \sigma^2$, use of the semivariance sharpens the original inequality.

If the distribution is known to be symmetric, then

$$\sigma_+^2 = \sigma_-^2 = \frac{1}{2}\sigma^2$$

and

$$\Pr(x \leq m - k\sigma) \leq \frac{1}{2k^2}.$$

This result agrees with that derived using standardised variables.

Note

The inequality with the lower semivariance has been found to be of use in estimating downside risk in finance and agriculture.^{[29][30][31]}

Selberg's inequality

Selberg derived an inequality for $P(x)$ when $a \leq x \leq b$.^[32] To simplify the notation let

$$Y = \alpha X + \beta$$

where

$$\alpha = \frac{2k}{b-a}$$

and

$$\beta = \frac{-(b+a)k}{b-a}.$$

The result of this linear transformation is to make $P(a \leq X \leq b)$ equal to $P(|Y| \leq k)$.

The mean (μ_X) and variance (σ_X) of X are related to the mean (μ_Y) and variance (σ_Y) of Y :

$$\mu_Y = \alpha\mu_X + \beta$$

$$\sigma_Y^2 = \alpha^2\sigma_X^2.$$

With this notation Selberg's inequality states that

$$P(|Y| < k) \geq \frac{(k - \mu_Y)^2}{(k - \mu_Y)^2 + \sigma_Y^2} \quad \text{if} \quad \sigma_Y^2 \leq \mu_Y(k - \mu_Y)$$

$$P(|Y| < k) \geq 1 - \frac{\sigma_Y^2 + \mu_Y^2}{k^2} \quad \text{if} \quad \mu_Y(k - \mu_Y) \leq \sigma_Y^2 \leq k^2 - \mu_Y^2$$

$$P(|Y| < k) \geq 0 \quad \text{if} \quad k^2 - \mu_Y^2 \leq \sigma_Y^2.$$

These are known to be the best possible bounds.^[33]

Cantelli's inequality

Cantelli's inequality^[34] due to Francesco Paolo Cantelli states that for a real random variable (X) with mean (μ) and variance (σ^2)

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

where $a \geq 0$.

This inequality can be used to prove a one tailed variant of Chebyshev's inequality with $k > 0$ ^[35]

$$\Pr(X - \mu \geq k\sigma) \leq \frac{1}{1 + k^2}.$$

The bound on the one tailed variant is known to be sharp. To see this consider the random variable X that takes the values

$$X = 1 \text{ with probability } \frac{\sigma^2}{1 + \sigma^2}$$

$$X = -\sigma^2 \text{ with probability } \frac{1}{1 + \sigma^2}.$$

Then $E(X) = 0$ and $E(X^2) = \sigma^2$ and $P(X < 1) = 1 / (1 + \sigma^2)$.

An application – distance between the mean and the median

The one-sided variant can be used to prove the proposition that for probability distributions having an expected value and a median, the mean and the median can never differ from each other by more than one standard deviation. To express this in symbols let μ , ν , and σ be respectively the mean, the median, and the standard deviation. Then

$$|\mu - \nu| \leq \sigma.$$

There is no need to assume that the variance is finite because this inequality is trivially true if the variance is infinite.

The proof is as follows. Setting $k = 1$ in the statement for the one-sided inequality gives:

$$\Pr(X - \mu \geq \sigma) \leq \frac{1}{2} \implies \Pr(X \geq \mu + \sigma) \leq \frac{1}{2}.$$

Changing the sign of X and of μ , we get

$$\Pr(X \leq \mu - \sigma) \leq \frac{1}{2}.$$

Thus the median is within one standard deviation of the mean.

A proof using Jensen's inequality also exists.

Bhattacharyya's inequality

Bhattacharyya^[36] extended Cantelli's inequality using the third and fourth moments of the distribution.

Let $\mu = 0$ and σ^2 be the variance. Let $\gamma = E(X^3) / \sigma^3$ and $\kappa = E(X^4) / \sigma^4$.

If $k^2 - k\gamma - 1 > 0$ then

$$P(X > k\sigma) \leq \frac{\kappa - \gamma^2 - 1}{(\kappa - \gamma^2 - 1)(1 + k^2) + (k^2 - k\gamma - 1)}.$$

The necessity of $k^2 - k\gamma - 1 > 0$ requires that k be reasonably large.

Mitzenmacher and Upfal's inequality

Mitzenmacher and Upfal^[37] note that

$$(X - \mathbb{E}[X])^{2k} > 0$$

for any real $k > 0$ and that

$$\mathbb{E}[(X - \mathbb{E}(X))^{2k}]$$

is the $2k^{\text{th}}$ central moment. They then show that for $t > 0$

$$\Pr\left(|X - \mathbb{E}[X]| > t \mathbb{E}[(X - \mathbb{E}[X])^{2k}]^{1/2k}\right) \leq \frac{1}{t^{2k}}.$$

For $k = 1$ we obtain Chebyshev's inequality. For $t \geq 1$, $k > 2$ and assuming that the k^{th} moment exists, this bound is tighter than Chebyshev's inequality.

Related inequalities

Several other related inequalities are also known.

Zelen's inequality

Zelen has shown that^[38]

$$\Pr(X - \mu \geq k\sigma) \leq \left[1 + k^2 + \frac{(k^2 - k\theta_3 - 1)^2}{\theta_4 - \theta_3^2 - 1}\right]^{-1}$$

with

$$k \geq \frac{\theta_3 + \sqrt{\theta_3^2 + 4}}{2}, \qquad \theta_m = \frac{M_m}{\sigma}$$

where M_m is the m -th moment and σ is the standard deviation.

He, Zhang and Zhang's inequality

For any collection of n non-negative independent random variables X_i with expectation 1 ^[39]

$$\Pr\left(\frac{\sum_{i=1}^n X_i}{n} - 1 \geq \frac{1}{n}\right) \leq \frac{7}{8}.$$

Hoeffding's lemma

Let X be a random variable with $a \leq X \leq b$ and $\mathbb{E}[X] = 0$, then for any $s > 0$, we have

$$E\left[e^{sX}\right] \leq e^{\frac{1}{8}s^2(b-a)^2}.$$

Van Zuijlen's bound

Let X_i be a set of independent Rademacher random variables: $\Pr(X_i = 1) = \Pr(X_i = -1) = 0.5$. Then^[40]

$$\Pr\left(\left|\frac{\sum_{i=1}^n X_i}{\sqrt{n}}\right| \leq 1\right) \geq 0.5.$$

The bound is sharp and better than that which can be derived from the normal distribution (approximately $\Pr > 0.31$).

Unimodal distributions

A distribution function F is unimodal at ν if its cumulative distribution function is convex on $(-\infty, \nu)$ and concave on (ν, ∞) .^[41] An empirical distribution can be tested for unimodality with the dip test.^[42]

In 1823 Gauss showed that for a unimodal distribution with a mode of zero^[43]

$$P(|X| \geq k) \leq \frac{4 E(X^2)}{9k^2} \quad \text{if} \quad k^2 \geq \frac{4}{3} E(X^2),$$

$$P(|X| \geq k) \leq 1 - \frac{k^2}{3 E(X^2)} \quad \text{if} \quad k^2 \leq \frac{4}{3} E(X^2).$$

If the mode (ν) is not zero and the mean (μ) and standard deviation (σ) are both finite then denoting the root mean square deviation from the mode by ω , we have

$$\sigma \leq \omega \leq 2\sigma,$$

and

$$|\nu - \mu| \leq \sqrt{\frac{3}{4}}\omega.$$

Winkler in 1866 extended Gauss' inequality to r^{th} moments ^[44] where $r > 0$ and the distribution is unimodal with a mode of zero:

$$P(|X| \geq k) \leq \left(\frac{r}{r+1}\right)^r \frac{E(|X|)^r}{k^r} \quad \text{if} \quad k^r \geq \frac{r^r}{(r+1)^{r+1}} E(|X|^r),$$

$$P(|X| \geq k) \leq \left(1 - \left[\frac{k^r}{(r+1) E(|X|^r)}\right]^{1/r}\right) \quad \text{if} \quad k^r \leq \frac{r^r}{(r+1)^{r+1}} E(|X|^r).$$

Gauss' bound has been subsequently sharpened and extended to apply to departures from the mean rather than the mode due to the Vysochanskiĭ–Petunin inequality.

The Vysochanskiĭ–Petunin inequality has been extended by Dharmadhikari and Joag-Dev^[45]

$$P(|X| > k) \leq \max\left(\left[\frac{r}{(r+1)k}\right]^r E|X^r|, \frac{s}{(s-1)k^r} E|X^r| - \frac{1}{s-1}\right)$$

where s is a constant satisfying both $s > r + 1$ and $s(s - r - 1) = r^r$ and $r > 0$.

It can be shown that these inequalities are the best possible and that further sharpening of the bounds requires that additional restrictions be placed on the distributions.

Unimodal symmetrical distributions

The bounds on this inequality can also be sharpened if the distribution is both unimodal and symmetrical.^[46] An empirical distribution can be tested for symmetry with a number of tests including McWilliam's R*.^[47] It is known that the variance of a unimodal symmetrical distribution with finite support $[a, b]$ is less than or equal to $(b - a)^2 / 12$.^[48]

Let the distribution be supported on the finite interval $[-N, N]$ and the variance be finite. Let the mode of the distribution be zero and rescale the variance to 1. Let $k > 0$ and assume $k < 2N/3$. Then^[46]

$$P(X \geq k) \leq \frac{1}{2} - \frac{k}{2\sqrt{3}} \quad \text{if} \quad 0 \leq k \leq \frac{2}{\sqrt{3}},$$

$$P(X \geq k) \leq \frac{2}{9k^2} \quad \text{if} \quad \frac{2}{\sqrt{3}} \leq k \leq \frac{2N}{3}.$$

If $0 < k \leq 2/\sqrt{3}$ the bounds are reached with the density^[46]

$$f(x) = \frac{1}{2\sqrt{3}} \quad \text{if} \quad |x| < \sqrt{3}$$

$$f(x) = 0 \quad \text{if} \quad |x| \geq \sqrt{3}.$$

If $2/\sqrt{3} < k \leq 2N/3$ the bounds are attained by the distribution

$$(1 - \beta_k)\delta_0(x) + \beta_k f_k(x),$$

where $\beta_k = 4 / 3k^2$, δ_0 is the Dirac delta function and where

$$f_k(x) = \frac{1}{3k} \quad \text{if} \quad |x| < \frac{3k}{2},$$

$$f_k(x) = 0 \quad \text{if} \quad |x| \geq \frac{3k}{2}.$$

The existence of these densities shows that the bounds are optimal. Since N is arbitrary these bounds apply to any value of N .

The Camp–Meidell's inequality is a related inequality.^[49] For an absolutely continuous unimodal and symmetrical distribution

$$P(|X - \mu| \geq k\sigma) \leq 1 - \frac{k}{\sqrt{3}} \quad \text{if} \quad k \leq \frac{2}{\sqrt{3}},$$

$$P(|X - \mu| \geq k\sigma) \leq \frac{4}{9k^2} \quad \text{if} \quad k > \frac{2}{\sqrt{3}}.$$

The second of these inequalities is the same as the Vysochanskiĭ–Petunin inequality.

DasGupta has shown that if the distribution is known to be normal^[50]

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{3k^2}.$$

Notes

Effects of symmetry and unimodality

Symmetry of the distribution decreases the inequality's bounds by a factor of 2 while unimodality sharpens the bounds by a factor of 4/9.

Because the mean and the mode in a unimodal distribution differ by at most $\sqrt{3}$ standard deviations^[51] at most 5% of a symmetrical unimodal distribution lies outside $(2\sqrt{10} + 3\sqrt{3})/3$ standard deviations of the mean (approximately 3.840 standard deviations). This is sharper than the bounds provided by the Chebyshev inequality (approximately 4.472 standard deviations).

These bounds on the mean are less sharp than those that can be derived from symmetry of the distribution alone which shows that at most 5% of the distribution lies outside approximately 3.162 standard deviations of the mean. The Vysochanskiĭ–Petunin inequality further sharpens this bound by showing that for such a distribution that at most 5% of the distribution lies outside $4\sqrt{5}/3$ (approximately 2.981) standard deviations of the mean.

Normal distributions

DasGupta's inequality states that for a normal distribution at least 95% lies within approximately 2.582 standard deviations of the mean. This is less sharp than the true figure (approximately 1.96 standard deviations of the mean).

Bounds for specific distributions

DasGupta has determined a set of best possible bounds for a normal distribution for this inequality.^[50]

Steliga and Szynal have extended these bounds to the Pareto distribution.^[8]

Zero means

When the mean (μ) is zero Chebyshev's inequality takes a simple form. Let σ^2 be the variance. Then

$$P(|X| \geq 1) \leq \min(1, \sigma^2).$$

With the same conditions Cantelli's inequality takes the form

$$P(X \geq 1) \leq \frac{\sigma^2}{1 + \sigma^2}.$$

Unit variance

If in addition $E(X^2) = 1$ and $E(X^4) = \psi$ then for any $0 \leq \epsilon \leq 1$ ^[52]

$$P(|X| > \epsilon) \geq \frac{(1 - \epsilon^2)^2}{\psi - 1 + (1 - \epsilon^2)^2} \geq \frac{(1 - \epsilon^2)^2}{\psi}.$$

The first inequality is sharp.

It is also known that for a random variable obeying the above conditions that^[53]

$$P(X \geq \epsilon) \geq \frac{C_0}{\psi} - \frac{C_1}{\sqrt{\psi}}\epsilon + \frac{C_2}{\psi\sqrt{\psi}}\epsilon$$

where

$$\begin{aligned} C_0 &= 2\sqrt{3} - 3 \quad (\approx 0.464), \\ C_1 &= 1.397, \\ C_2 &= 0.0231. \end{aligned}$$

It is also known that^[53]

$$P(X > 0) \geq \frac{C_0}{\psi}.$$

The value of C₀ is optimal and the bounds are sharp if

$$\psi \geq \frac{3}{\sqrt{3} + 1} \quad (\approx 1.098).$$

If

$$\psi \leq \frac{3}{\sqrt{3} + 1}$$

then the sharp bound is

$$P(X > 0) \geq \frac{2}{3 + \psi + \sqrt{(1 + \psi)^2 - 4}}.$$

Integral Chebyshev inequality

There is a second (less well known) inequality also named after Chebyshev^[54]

If *f*, *g* : [*a*, *b*] → **R** are two monotonic functions of the same monotonicity, then

$$\frac{1}{b-a} \int_a^b f(x)g(x) \, dx \geq \left[\frac{1}{b-a} \int_a^b f(x) \, dx \right] \left[\frac{1}{b-a} \int_a^b g(x) \, dx \right].$$

If *f* and *g* are of opposite monotonicity, then the above inequality works in the reverse way.

This inequality is related to Jensen's inequality,^[55] Kantorovich's inequality,^[56] the Hermite–Hadamard inequality^[56] and Walter's conjecture.^[57]

Other inequalities

There are also a number of other inequalities associated with Chebyshev:

- Chebyshev's sum inequality
- Chebyshev–Markov–Stieltjes inequalities

Haldane's transformation

One use of Chebyshev's inequality in applications is to create confidence intervals for variates with an unknown distribution. Haldane noted,^[58] using an equation derived by Kendall,^[59] that if a variate (*x*) has a zero mean, unit variance and both finite skewness (*γ*) and kurtosis (*κ*) then the variate can be converted to a normally distributed standard score (*z*):

$$z = x - \frac{\gamma}{6}(x^2 - 1) + \frac{x}{72}[2\gamma^2(4x^2 - 7) - 3\kappa(x^2 - 3)] + \cdots$$

This transformation may be useful as an alternative to Chebyshev's inequality or as an adjunct to it for deriving confidence intervals for variates with unknown distributions.

While this transformation may be useful for moderately skewed and/or kurtotic distributions, it performs poorly when the distribution is markedly skewed and/or kurtotic.

Notes

The Environmental Protection Agency has suggested best practices for the use of Chebyshev's inequality for estimating confidence intervals.^[60] This caution appears to be justified as its use in this context may be seriously misleading.^[61]

See also

- Multidimensional Chebyshev's inequality
- Concentration inequality - a summary of tail-bounds on random variables.
- Cornish–Fisher expansion
- Eaton's inequality
- Kolmogorov's inequality
- Proof of the weak law of large numbers using Chebyshev's inequality
- Le Cam's theorem
- Paley–Zygmund inequality
- Vysochanskiĭ–Petunin inequality — a stronger result applicable to unimodal probability distributions

References

- Kvanli, Alan H.; Pavur, Robert J.; Keeling, Kellie B. (2006). *Concise Managerial Statistics*. cEngage Learning. pp. 81–82. ISBN 9780324223880.
- Chernick, Michael R. (2011). *The Essentials of Biostatistics for Physicians, Nurses, and Clinicians*. John Wiley & Sons. pp. 49–50. ISBN 9780470641859.
- Knuth, Donald (1997). *The Art of Computer Programming: Fundamental Algorithms, Volume 1* (3rd ed.). Reading, Massachusetts: Addison–Wesley. ISBN 0-201-89683-4. Retrieved 1 October 2012.
- Bienaymé I.-J. (1853) Considérations à l'appui de la découverte de Laplace. Comptes Rendus de l'Académie des Sciences 37: 309–324
- Tchebichef, P. (1867). "Des valeurs moyennes". *Journal de mathématiques pures et appliquées*. 2 12: 177–184.
- Markov A. (1884) On certain applications of algebraic continued fractions, Ph.D. thesis, St. Petersburg
- Grafakos, Lukas (2004). *Classical and Modern Fourier Analysis*. Pearson Education Inc. p. 5.
- Steliga, Katarzyna; Szynd, Dominik (2010). "On Markov-Type Inequalities" (PDF). *International Journal of Pure and Applied Mathematics* 58 (2): 137–152. ISSN 1311-8080. Retrieved 10 October 2012.
- Ferentinos, K (1982). "On Tchebycheff type inequalities". *Trabajos Estadist Investigacion Oper* 33: 125–132.
- Berge P. O. (1938) A note on a form of Tchebycheff's theorem for two variables. Biometrika 29, 405–406
- Lal D. N. (1955) A note on a form of Tchebycheff's inequality for two or more variables. Sankhya 15(3):317–320
- Isii K. (1959) On a method for generalizations of Tchebycheff's inequality. Ann Inst Stat Math 10: 65–88
- Birnbaum, Z. W.; Raymond, J.; Zuckerman, H. S. (1947). "A Generalization of Tshebyshev's Inequality to Two Dimensions". *The Annals of Mathematical Statistics* 18 (1): 70–79. doi:10.1214/aoms/1177730493. ISSN 0003-4851. MR 19849. Zbl 0032.03402. Retrieved 7 October 2012.
- Kotz, Samuel; Balakrishnan, N.; Johnson, Norman L. (2000). *Continuous Multivariate Distributions, Volume 1, Models and Applications* (2nd ed.). Boston [u.a.]: Houghton Mifflin. ISBN 978-0-471-18387-7. Retrieved 7 October 2012.
- Olkin, Ingram; Pratt, John W. (1958). "A Multivariate Tchebycheff Inequality". *The Annals of Mathematical Statistics* 29 (1): 226–234. doi:10.1214/aoms/1177706720. MR MR93865. Zbl 0085.35204. Retrieved 2 October 2012.
- Godwin H. J. (1964) Inequalities on distribution functions. New York, Hafner Pub. Co.
- Xinjia Chen (2007). "A New Generalization of Chebyshev Inequality for Random Vectors". arXiv:0707.0805v2.
- Vakhania, Nikolai Nikolaevich. Probability distributions on linear spaces. New York: North Holland, 1981.
- Section 2.1 (<http://www.math.utah.edu/~firas/Papers/rassoul-seppalainen-ldp.pdf>)
- Baranoski, Gladimir V. G.; Rokne, Jon G.; Xu, Guangwu (15 May 2001). "Applying the exponential Chebyshev inequality to the nondeterministic computation of form factors". *Journal of Quantitative Spectroscopy and Radiative Transfer* 69 (4): 199–200. Bibcode:2001JQSRT..69..447B. doi:10.1016/S0022-4073(00)00095-9. Retrieved 2 October 2012. (the references for this article are corrected by Baranoski, Gladimir V. G.; Rokne, Jon G.; Guangwu Xu (15 January 2002). "Corrigendum to: 'Applying the exponential Chebyshev inequality to the nondeterministic computation of form factors'". *Journal of Quantitative Spectroscopy and Radiative Transfer* 72 (2): 199–200. Bibcode:2002JQSRT..72..199B. doi:10.1016/S0022-4073(01)00171-6. Retrieved 2 October 2012.)
- Dufour (2003) Properties of moments of random variables (http://www2.cirano.qc.ca/~dufourj/Web_Site/ResE/Dufour_1999_C_TS_Moments.pdf)
- Niemitalo O. (2012) One-sided Chebyshev-type inequalities for bounded probability distributions. (<http://yehar.com/blog/?p=1225>)
- Saw, John G.; Yang, Mark C. K.; Mo, Tse Chin (1984). "Chebyshev Inequality with Estimated Mean and Variance". *The American Statistician* (American Statistical Association) 38 (2): 130–2. doi:10.2307/2683249. ISSN 0003-1305. JSTOR 2683249 – via JSTOR. (registration required (help)).
- Konijn, Hendrik S. (February 1987). "Distribution-Free and Other Prediction Intervals". *The American Statistician* (American Statistical Association) 41 (1): 11–15. doi:10.2307/2684311. JSTOR 2684311 – via JSTOR. (registration required (help)).
- Kabán, Ata (2012). "Non-parametric detection of meaningless distances in high dimensional data". *Statistics and Computing* 22 (2): 375–85. doi:10.1007/s11222-011-9229-0.
- Beasley, T. Mark; Page, Grier P.; Brand, Jaap P. L.; Gadbury, Gary L.; Mountz, John D.; Allison, David B. (January 2004). "Chebyshev's inequality for nonparametric testing with small *N* and α in microarray research". *Journal of the Royal Statistical Society. C (Applied Statistics)* 53 (1): 95–108. doi:10.1111/j.1467-9876.2004.00428.x. ISSN 1467-9876. Retrieved 3 October 2012.
- Savage, I. Richard. "Probability inequalities of the Tchebycheff type." Journal of Research of the National Bureau of Standards-B. Mathematics and Mathematical Physics B 65 (1961): 211-222 (http://nvlpubs.nist.gov/nistpubs/jres/65B/jresv65Bn3p211_A1b.pdf)
- Ion, Roxana Alice (2001). "Chapter 4: Sharp Chebyshev-type inequalities". *Nonparametric Statistical Process Control*. Universiteit van Amsterdam. ISBN 9057760762. Retrieved 1 October 2012.
- Berck, Peter; Hihn, Jairus M. (May 1982). "Using the Semivariance to Estimate Safety-First Rules" (PDF). *American Journal of Agricultural Economics* (Oxford University Press) 64 (2): 298–300. doi:10.2307/1241139. ISSN 0002-9092. Retrieved 8 October 2012.
- Nantell, Timothy J.; Price, Barbara (June 1979). "An Analytical Comparison of Variance and Semivariance Capital Market Theories". *The Journal of Financial and Quantitative Analysis* (University of Washington School of Business Administration) 14 (2): 221–42. doi:10.2307/2330500. JSTOR 2330500 – via JSTOR. (registration required (help)).
- Neave, Edwin H.; Ross, Michael N.; Yang, Jun (2009). "Distinguishing upside potential from downside risk". *Management Research News* 32 (1): 26–36. doi:10.1108/01409170910922005. ISSN 0140-9174. (registration required (help)).
- Selberg, Henrik L. (1940). "Zwei Ungleichungen zur Ergänzung des Tchebycheffschen Lemmas" [Two Inequalities Supplementing the Tchebycheff Lemma]. *Skandinavisk Aktuarietidskrift (Scandinavian Actuarial Journal)* (in German) 1940 (3–4): 121–125. doi:10.1080/03461238.1940.10404804. ISSN 0346-1238. OCLC 610399869. Retrieved 7 October 2012.
- Conlon, J.; Dulá, J. H. "A geometric derivation and interpretation of Tchebyscheff's Inequality" (PDF). Retrieved 2 October 2012.
- Cantelli F. (1910) Intorno ad un teorema fondamentale della teoria del rischio. Bolletino dell'Associazione degli Attuari Italiani
- Grimmett and Stirzaker, problem 7.11.9. Several proofs of this result can be found in Chebyshev's Inequalities (<http://www.mcdowella.demon.co.uk/Chebyshev.html>) by A. G. McDowell.
- Bhattacharyya, B. B. (1987). "One-sided chebyshev inequality when the first four moments are known". *Communications in Statistics – Theory and Methods* 16 (9): 2789–91. doi:10.1080/03610928708829540. ISSN 0361-0926.
- Mitzenmacher, Michael; Upfal, Eli (January 2005). *Probability and Computing: Randomized Algorithms and Probabilistic Analysis* (Repr. ed.). Cambridge [u.a.]: Cambridge Univ. Press. ISBN 9780521835404. Retrieved 6 October 2012.
- Zelen M. (1954) Bounds on a distribution function that are functions of moments to order four. J Res Nat Bur Stand 53:377–381
- He, S.; Zhang, J.; Zhang, S. (2010). "Bounding probability of small deviation: A fourth moment approach". *Mathematics of operations research* 35 (1): 208–232. doi:10.1287/moor.1090.0438.
- Martien C. A. van Zuijlen (2011) On a conjecture concerning the sum of independent Rademacher random variables (<http://arxiv.org/abs/1112.4988>)
- Feller, William (1966). *An Introduction to Probability Theory and Its Applications, Volume 2* (2 ed.). Wiley. p. 155. Retrieved 6 October 2012.
- Hartigan J. A., Hartigan P. M. (1985) "The dip test of unimodality" (<http://projecteuclid.org/euclid.aos/1176346577>). *Annals of Statistics* 13(1):70–84 doi:10.1214/aos/1176346577 (<https://dx.doi.org/10.1214/aos/1176346577>) MR 773153 (<http://www.ams.org/mathscinet-getitem?mr=773153>)
- Gauss C. F. Theoria Combinationis Observationum Erroribus Minimis Obnoxiae. Pars Prior. Pars Posterior. Supplementum. Theory of the Combination of Observations Least Subject to Errors. Part One. Part Two. Supplement. 1995. Translated by G. W. Stewart. Classics in Applied Mathematics Series, Society for Industrial and Applied Mathematics, Philadelphia
- Winkler A. (1886) Math-Natur theorie Kl. Akad. Wiss Wien Zweite Abt 53, 6–41

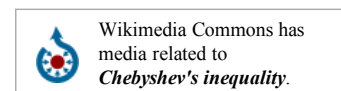
45. Dharmadhikari, S. W.; Joag-Dev, K. (1985). "The Gauss–Tchebyshev inequality for unimodal distributions". *Teor Veroyatnost i Primenen* **30** (4): 817–820.
46. Clarkson, Eric; Denny, J. L.; Shepp, Larry (2009). "ROC and the bounds on tail probabilities via theorems of Dubins and F. Riesz". *The Annals of Applied Probability* **19** (1): 467–76. doi:10.1214/08-AAP536. PMC 2828638. PMID 20191100.
47. McWilliams, Thomas P. (1990). "A Distribution-Free Test for Symmetry Based on a Runs Statistic". *Journal of the American Statistical Association* (American Statistical Association) **85** (412): 1130–3. doi:10.2307/2289611. ISSN 0162-1459. JSTOR 2289611 – via JSTOR. (registration required (help)).
48. Seaman, John W., Jr.; Young, Dean M.; Odell, Patrick L. (1987). "Improving small sample variance estimators for bounded random variables". *Industrial Mathematics* **37**: 65–75. ISSN 0019-8528. Zbl 0637.62024.
49. Bickel, Peter J.; Krieger, Abba M. (1992). "Extensions of Chebyshev's Inequality with Applications". *Probability and Mathematical Statistics* (Wy dawnictwo Uniwersytetu Wrocławskiego) **13** (2): 293–310. ISSN 0208-4147. Retrieved 6 October 2012.
50. DasGupta A. (2000) Best constants in Chebychev inequalities with various applications. *Metrika* 5(1):185–200
51. "More thoughts on a one tailed version of Chebyshev's inequality – by Henry Bottomley". se16.info. Retrieved 2012-06-12.
52. Godwin H. J. (1964) Inequalities on distribution functions. (Chapter 3) New York, Hafner Pub. Co.
53. Lesley F. D., Rotar V. I. (2003) Some remarks on lower bounds of Chebyshev's type for half-lines. *J Inequalities Pure Appl Math* 4(5) Art 96
54. Fink, A. M.; Jodeit, Max, Jr. (1984). Tong, Y. L.; Gupta, Shanti S., eds. "On Chebyshev's Other Inequality". *Institute of Mathematical Statistics Lecture Notes – Monograph Series*. Inequalities in Statistics and Probability: Proceedings of the Symposium on Inequalities in Statistics and Probability, October 27–30, 1982, Lincoln, Nebraska **5**: 115–120. doi:10.1214/lnms/1215465637. ISBN 0-940600-04-8. MR 789242. Retrieved 7 October 2012.
55. Niculescu, Constantin P.; Pečarić, Josip (2010). "An extension of Chebyshev's inequality and its connection with Jensen's inequality". *Journal of Inequalities and Applications* **6** (4): 451–462. doi:10.1155/S1025583401000273. ISSN 1025-5834. Retrieved 6 October 2012.
56. Niculescu, Constantin P.; Pečarić, Josip (2010). "The Equivalence of Chebyshev's Inequality to the Hermite–Hadamard Inequality" (PDF). *Mathematical Reports* (Publishing House of the Romanian Academy) **12** (62): 145–156. ISSN 1582-3067. Retrieved 6 October 2012.
57. Malamud, S. M. (15 February 2001). "Some complements to the Jensen and Chebyshev inequalities and a problem of W. Walter". *Proceedings of the American Mathematical Society* **129** (9): 2671–2678. doi:10.1090/S0002-9939-01-05849-X. ISSN 0002-9939. MR 1838791. Retrieved 7 October 2012.
58. Haldane, J. B. (1952). "Simple tests for bimodality and bitangentiality". *Annals of Eugenics* **16** (4): 359–364. doi:10.1111/j.1469-1809.1951.tb02488.x.
59. Kendall M. G. (1943) The Advanced Theory of Statistics, 1. London
60. Calculating Upper Confidence Limits for Exposure Point Concentrations at hazardous Waste Sites (PDF) (Report). Office of Emergency and Remedial Response of the U.S. Environmental Protection Agency. December 2002. Retrieved 1 October 2012.
61. "Statistical Tests: The Chebyshev UCL Proposal". *Quantitative Decisions*. 25 March 2001. Retrieved 26 November 2015.

Further reading

- A. Papoulis (1991), *Probability, Random Variables, and Stochastic Processes*, 3rd ed. McGraw–Hill. ISBN 0-07-100870-5. pp. 113–114.
- G. Grimmett and D. Stirzaker (2001), *Probability and Random Processes*, 3rd ed. Oxford. ISBN 0-19-857222-0. Section 7.3.

External links

- Hazewinkel, Michiel, ed. (2001), "Chebyshev inequality in probability theory", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Formal proof (http://mws.cs.ru.nl/mwiki/random_2.html#T7) in the Mizar system.



Retrieved from "https://en.wikipedia.org/w/index.php?title=Chebyshev%27s_inequality&oldid=707382982"

Categories: Probabilistic inequalities | Probability theory | Statistical inequalities

- This page was last modified on 28 February 2016, at 15:17.
- Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.