The Generalized Mean and Error Analysis

The Affiliated High School of South China Normal University

Author: Ketian Zhang

Faculty Adviser: Yiwen Huang

The Generalized Mean and Error Analysis

Abstract: In this paper, the definition of generalized mean $\overline{x_f} = f^{-1}[\frac{1}{n}\sum_{i=1}^{n} f(x_i)]$ is introduced. It summarizes the common characterizations and also expands the meaning of arithmetic, geometric and harmonic means. Secondly, the definition of average error is introduced, and the method of reducing error by averaging the values out is discussed. Based on the properties of convex function and the Jensen Inequality, the judgment of $\overline{x_f} \le \overline{x_g}$ and the necessary and sufficient condition of $\overline{x_f} = \overline{x_g}$ are discussed. Finally, the definition of effect coefficient which measures the effect of extreme values on a generalized mean is introduced. Some applications are contained in the paper as well.

Keywords: Mean; Error; Convex function; Jensen Inequality.

The aim and background of the research: What are the similarities of arithmetic, geometric and harmonic means? Whether the meaning of mean could be expanded? In scientific experiments, we often reduce errors by averaging the values. Are the effects different if we choose different means? In order to reduce the effect of extreme values, how to choose a kind of mean?

【Main results】

The notations used in this paper are listed as follows.

I. Let $U_1, U_2, ..., U_n$ and U be continuous random variables which satisfy (0，1]-uniform distribution.

II. Denote by $A_n = \frac{1}{n}\sum_{i=1}^{n} U_i$, $G_n = \left(\prod_{i=1}^{n} Ui\right)^{\frac{1}{n}}$, $H_n = \frac{n}{\sum_{i=1}^{n}\frac{1}{Ui}}$.

III. For a continuous random variable X, let $F_X(x)$ and $p_X(x)$ denote the distribution function and the density function of X, respectively.

1. Generalized mean

Definition 1. If the inverse function of $f(x)$ exists, then $\overline{x_f} = f^{-1}[\frac{1}{n}\sum_{i=1}^{n} f(x_i)]$ is called the $f(x)$-mean of $x_1, x_2, \ldots, x_n$.

Theorem 1. The arithmetic mean is x-mean, the geometric mean is $\ln x$-mean, and the harmonic mean is $\frac{1}{x}$-mean.

Proof. (1) Let $a(x) = x$. Then $a^{-1}(x) = x \Rightarrow \overline{x_a} = \frac{1}{n}\sum_{i=1}^{n} x_i$, that is, the arithmetic mean of $x_1, x_2, \ldots, x_n$.

(2) Let $g(x) = \ln x$. Then $g^{-1}(x) = e^x$. For $x_1, x_2, \ldots, x_n \in (0, +\infty)$,

$$\overline{x_g} = e\text{^}(\frac{1}{n}\sum_{i=1}^{n}\ln x_i) = e\text{^}(\frac{1}{n}\ln\prod_{i=1}^{n} x_i) = e\text{^}(\ln\sqrt[n]{\prod_{i=1}^{n} x_i}) = \sqrt[n]{\prod_{i=1}^{n} x_i},$$

which is the geometric mean of $x_1, x_2, \ldots, x_n$.

(3) Let $h(x) = \frac{1}{x}$. Then $h^{-1}(x) = \frac{1}{x}$. For $x_1, x_2, \ldots, x_n \in (0, +\infty)$,

$$\overline{x_h} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}},$$

which is the harmonic mean of $x_1, x_2, \ldots, x_n$. ■

Theorem 2. If the inverse function of $f(x)$ exists, and $g(x) = af(x) + b$ ($a \neq 0$), then $\overline{x_g} \equiv \overline{x_f}$.

Proof. Note that the inverse function of $g(x)$ is $g^{-1}(x) = f^{-1}(\frac{x-b}{a})$. Then

$$\overline{x_g} = g^{-1}\{\frac{1}{n}\sum_{i=1}^{n}[af(x_i)+b]\}$$

$$= g^{-1}[a\frac{1}{n}\sum_{i=1}^{n} f(x_i)+b]$$

$$= f^{-1}[\frac{1}{n}\sum_{i=1}^{n} f(x_i)] = \overline{x_f},$$

that is, $\overline{x_g} = \overline{x_f}$. ■

2. Average error

The variance and standard deviation usually used in mathematical statistics can characterize the data volatility, but the meaning of their specific value is not so straightforward. Therefore, the definition of average error is introduced. Moreover, we shall study some problems by using this definition.

Definition 2. For a random variable X, $\overline{e_X} = E(|X - E(X)|)$ is defined to be the average error of X. Moreover, $\bar{e}(\overline{x_f}) = E(|\overline{x_f} - E(X)|)$ is called the $f(x)$-average error of X.

2.1. Basic properties of average error

Obviously, for a continuous random variable X, $\overline{e_X} = \int_{-\infty}^{+\infty} |X - E(X)| p_X(x)dx$.

Theorem 3. If the standard deviation of a random variable X exists, then

$$\overline{e_X} \leq \sigma(X).$$

Proof. If X is a continuous random variable, then $\int_{-\infty}^{+\infty} p_X(x)dx = 1$.

Since x^2 is a convex function, then by the Jensen Inequality, we have

$$[\int_{-\infty}^{+\infty}|X-E(X)|p_X(x)dx]^2 \le \int_{-\infty}^{+\infty}[X-E(X)]^2 p_X(x)dx,$$

that is, $$\overline{e_x}^2 \le Var(X).$$

Hence $$\overline{e_X} \le \sigma(X).$$

If X is a discrete random variable, the result can be proved similarly. ■

Theorem 4. If the mathematical expectation and average error of a random variable X both exist, then for any constant ε> 0, we have

$$P(|X-E(X)| \ge \varepsilon) \le \frac{\overline{e_X}}{\varepsilon}.$$

Proof. If X is a continuous random variable, then

$$P(|X-E(X)| \ge \varepsilon) = \int_{\{x;|X-E(X)|\ge\varepsilon\}} p_X(x)dx \le \int_{\{x;|X-E(X)|\ge\varepsilon\}} \frac{|X-E(X)|}{\varepsilon} p_X(x)dx$$

$$\le \frac{1}{\varepsilon}\int_{-\infty}^{+\infty}|X-E(X)|p_X(x)dx = \frac{\overline{e_X}}{\varepsilon}.$$

If X is a discrete random variable, then the result can be proved similarly. ■

The result of Theorem 4 is similar to Chebyshev's Inequality. They both give the upper bound of large deviation probability of occurrence. Since the probability of any events is no more than 1, we call "the above upper bound is less than 1" "meaningful". For Theorem 4, when $\varepsilon > \overline{e_X}$, the upper bound is meaningful; for Chebyshev's Inequality, when $\frac{Var(X)}{\varepsilon^2} < 1$ (namely, $\varepsilon > \sigma(X)$), the upper bound is meaningful. If follows from $\overline{e_X} \le \sigma(X)$ that the meaningful scope of Theorem 4 is large than that of Chebyshev's Inequality.

Example 1. The average error of U.
Solution. Since

$$p_U(x) = \begin{cases} 0 & (x\leqslant 0 \text{ or } x>1), \\ 1 & (0 < x \le 1), \end{cases}$$

and $E(U) = \frac{1}{2}$, hence

$$\overline{e_U} = \int_0^1 \left|x-\frac{1}{2}\right| \cdot 1dx = \int_0^{\frac{1}{2}}(\frac{1}{2}-x)dx + \int_{\frac{1}{2}}^1 (x-\frac{1}{2})dx = \frac{1}{4}.$$

Answer. The average error of U is $\frac{1}{4}$.

Example 2. Let $X \sim N(\mu, \sigma^2)$. Then $\overline{e_X} = \sqrt{\frac{2}{\pi}}\sigma$.

Proof.

$$\overline{e_X} = \int_{-\infty}^{+\infty} |x-\mu| \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \int_{-\infty}^{+\infty} |x| \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= 2\int_{0}^{+\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx.$$

Let $t = \frac{x}{\sqrt{2}\sigma}$. Then

$$\overline{e_X} = \frac{2}{\sqrt{\pi}} \int_0^{+\infty} te^{-t^2} dt \cdot \sqrt{2}\sigma = (-\frac{1}{2}e^{-t^2})\Big|_0^{+\infty} \cdot 2\sqrt{\frac{2}{\pi}}\sigma = \sqrt{\frac{2}{\pi}}\sigma. \quad ■$$

Theorem 5. Let $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ and $a(x) = x$. Then $\bar{e}(\overline{X_a}) = \sqrt{\frac{2}{n\pi}}\sigma$.

Proof. By the additivity of normal distribution, $\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$. Then

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n}), \text{ and } E(\overline{X}) = E(X_i) \quad (i = 1,2,...,n).$$

Hence

$$\bar{e}(\overline{X}) = \overline{e\overline{X}} = \sqrt{\frac{2}{\pi}} \cdot \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{2}{n\pi}}\sigma. \quad ■$$

It can be seen that the average error of X_i ($i = 1,2,...,n$) reduced $\sqrt{n}$-times by solving $\overline{X}$.

Example 3. The accuracy of an analytical balance is 5mg. But the accuracy in a chemical analysis need to be 1mg. Suppose the measured value satisfies normal distribution. How many times do the parallel measures need and then obtain the mean of *x* satisfying the requirement?

Solution. Suppose we need *n* times the parallel measures. From the above example, the error reduce 5-times implies $\sqrt{n}=5$, and it follows that $n=25$.

Theorem 6. Let $\{X_n\}$ be random variables with independent and identical distribution. Then $\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}X_i=E(X_i)$.

Proof. Let $E(X_i)=\mu$ and $Var(X_i)=\sigma^2$. Since $X_1, X_2, \ldots, X_n$ are independent and identical distribution, by Lindberg-Levy Central Limit Theorem,

$$\lim_{n\to+\infty}\frac{\sum_{i=1}^{n}X_i-n\mu}{\sigma\sqrt{n}}\sim N(0,1)\text{, and}$$

$$\overline{X}=\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}X_i\sim N(\mu,\frac{\sigma^2}{n}).$$

Since $$Var(\overline{X})=\lim_{n\to+\infty}\frac{\sigma^2}{n}=0,$$

$\overline{X}=\mu$, that is, $\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}X_i=E(X_i)$. ■

It shows that the number of parallel measures tends to positive infinity, the arithmetic mean of X_i tends to $E(X_i)$. For most surveying instruments with uniform scale, $E(X_i)$ is the true value.

Let $\{X_n\}$ be random variables with independent and identical distribution. Let $\overline{X_f}$ be the mean of $X_1, X_2, \cdots, X_n$. By Theorem 4, $\lim_{n\to+\infty}\overline{X_f}=f^{-1}[E(f(X_i))]$.

For example, if X_i satisfies (0，1]-uniform distribution, then $f(x)=\ln x$. It follows that $\lim_{n\to+\infty}\overline{X_f}=e\wedge(\int_0^1\ln x\cdot 1dx)$.

Since $$\int_0^1\ln x dx=(x\ln x-x)\Big|_0^1=-1-\lim_{x\to 0^+}(x\ln x-x)=-1,$$

$$\lim_{n\to+\infty}\overline{X_f}=\frac{1}{e}\neq E(X_i)=\frac{1}{2}.$$

It can be seen that if the number of parallel measures tends to positive infinity, then the $f(x)$-mean of X_i is not always equal to the true value.

Theorem 7. If $Y = aX + b$, then $\overline{e_Y} = |a|\overline{e_X}$, where *a* and *b* are constants.

Proof. Since $E(Y) = E(aX + b) = aE(X) + b$, we have

$$\overline{e_Y} = \int_{-\infty}^{+\infty} |(aX + b) - [aE(x) + b]| p_X(x)dx$$
$$= |a| \int_{-\infty}^{+\infty} |X - E(X)| p_X(x)dx$$
$$= |a|\overline{e_X}.$$

If the variation of $f(x)$ over $\{x | p_X(x) \neq 0\}$ is small, then for the random variable $Y = f(X)$, we have $Y \approx f'(x_0)X + b$ ($x_0 \in \{x | p_X(x) \neq 0\}$, b is a constant). By Theorem 5, $\overline{e_Y} \approx |f'(x_0)|\overline{e_X}$. ■

Example 4. Let the hydrogen ion concentration of a bottle of dilute hydrochloric acid be 0.056mol/L (the error is no larger than 0.001mol/L). Question: How many decimal places should the PH value of such dilute hydrochloric acid retain?

Solution. Suppose the hydrogen ion concentration is *x*, then $pH = -\lg x$ and $pH_x' = -\frac{1}{x \ln 10}$. Since $x \in [0.055, 0.057]$, then $-7.896 \leq pH_x' \leq -7.619$, the variation of pH_x' is relatively small. Hence $\overline{e_{pH}} \approx 7.7\overline{e_x}$, and the pH value retains two decimal places, that is,

$$pH = -\lg 0.056 = 1.25.$$

2.2. The average errors of A_2, G_2, H_2

2.2.1. The properties of A_2

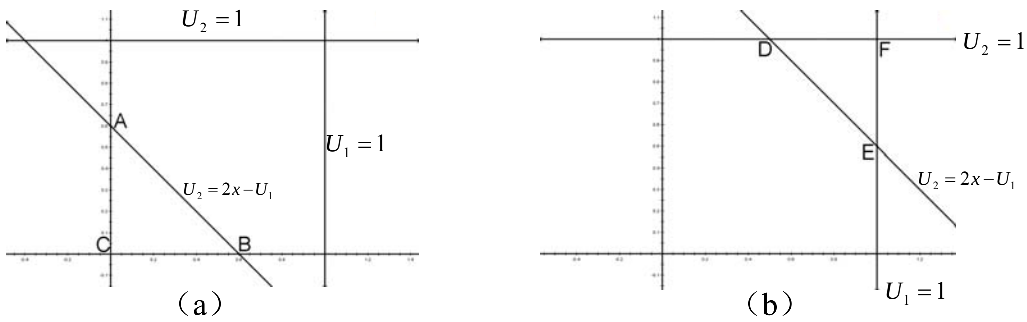
$$F_{A_2} = (\frac{U_1 + U_2}{2} \leq x) = P(U_2 \leq 2x - U_1).$$

Establish the plane rectangular coordinate system with abscissa axis U_1 and vertical axis U_2. Note that (U_1, U_2) locate in the field surrounded by $U_1 = 1$, $U_2 = 1$ and abscissa axis, vertical axis (except the origin) is an equally likely event. Then it satisfies the condition of geometric probability model.

See Fig. 1. (a), when $0 < x \leq \frac{1}{2}$, $P(U_2 \leq 2x - U_1) = S_{\Delta ABC}/1 = \frac{(2x)^2}{2} = 2x^2$.

See Fig. 1. (b), when $\frac{1}{2} < x < 1$,

$$P(U_2 \le 2x - U_1) = (1 - S_{\Delta DEF})/1 = 1 - \frac{(2-2x)^2}{2} = 1 - 2(1-x)^2.$$



（a） （b）

Figure 1.[1]

Since $$0 < \frac{U_1 + U_2}{2} \le 1,$$

then
$$F_{A2}(x) = \begin{cases} 0 & (x \le 0); \\ 2x^2 & (0 < x \le \frac{1}{2}); \\ 1 - 2(1-x)^2 & (\frac{1}{2} < x < 1); \\ 1 & (x \ge 1). \end{cases}$$

$$F_{A_2}(E(U)) = F_{A_2}(\frac{1}{2}) = \frac{1}{2}.$$

$$p_{A2}(x) = \begin{cases} 0 & (x \le 0 \text{ or } x \ge 1); \\ 4x & (0 < x \le \frac{1}{2}); \\ 4 - 4x & (\frac{1}{2} < x < 1). \end{cases}$$

$$\bar{e}(A_2) = \int_{-\infty}^{+\infty} |x - E(x)| p_{A2}(x) dx$$

$$= \int_0^{\frac{1}{2}} (\frac{1}{2} - x) \cdot 4x dx + \int_{\frac{1}{2}}^{1} (x - \frac{1}{2})(4 - 4x) dx$$

$$= (-\frac{4}{3}x^3 + x^2)\Big|_0^{\frac{1}{2}} + (-\frac{4}{3}x^3 + 3x^2 - 2x)\Big|_{\frac{1}{2}}^{1} = \frac{1}{6}.$$

2.2.2. The properties of G_2

$$F_{G2}(x) = P(\sqrt{U_1 U_2} \le x) = P(U_2 \le \frac{x^2}{U_1}).$$

[1] The density functions of the random variables corresponding to the vertices in the figure of the function may be equal to zero. Such vertices can not be attained, and the same below.

When $0<x<1$, we establish the plane rectangular coordinate system with abscissa axis U_1 and vertical axis U_2 (see Fig. 2.). Note that (U_1, U_2) locate in the field surrounded by $U_1=1$, $U_2=1$ and abscissa axis, vertical axis (except the origin) is an equally likely event. Then it satisfies the condition of geometric probability model. Since $U_2 \le \frac{x^2}{U_1}$ is the field surrounded by $U_2=\frac{x}{U_1}$, $U_1=1$, $U_2=1$ and abscissa axis, vertical axis, then $U_2=\frac{x}{U_1}$ intersects with $y=1$ at the intersection point $(x^2, 1)$. Hence

$$P(U_2 \le \frac{x^2}{U_1}) = (\int_{x^2}^{1} \frac{x^2}{U_1} dU_1 + x^2)/1$$

$$= x^2 \ln U_1 \Big|_{x^2}^{1} + x^2$$

$$= x^2 - x^2 \ln x^2 .$$

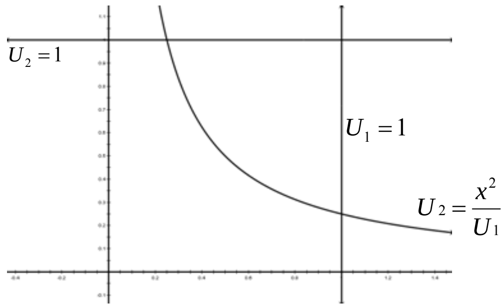


Figure 2.

Since $0<G_2\le 1$, then

$$F_{G_2}(x) = \begin{cases} 0 \quad (x \le 0); \\ x^2 - 2x^2 \ln x \quad (0<x<1); \\ 1 \quad (x \ge 1); \end{cases}$$

and $$F_{G_2}(E(U)) = F_{G_2}(\frac{1}{2}) = \frac{1}{4} + \ln\sqrt{2} = 0.5966,$$

$$p_{G_2}(x) = F_{G_2}{}'(x) = \begin{cases} 0 \quad (x \le 0 \quad or \quad x \ge 1); \\ -4x \ln x \quad (0<x<1); \end{cases}$$

$$\bar{e}(G_2) = \int_0^1 \left| x - \frac{1}{2} \right| (-4x \ln x) dx .$$

Let $g(x) = \frac{4}{9}x^3 + (x^2 - \frac{4}{3}x^3)\ln x - \frac{x^2}{2}$. It can be checked that

$$g'(x) = (x - \frac{1}{2})(-4x \ln x) .$$

Hence $$\bar{e}(G_2) = [g(1) - g(\frac{1}{2})] - [g(\frac{1}{2}) - \lim_{\mu \to 0} g(\mu)] = 0.1989 .$$

2.2.3. The properties of H_2

$F_{H_2}(x) = P(\frac{2}{\frac{1}{U_1}+\frac{1}{U_2}} \le x) = P(U_2 \le \frac{xU_1}{2U_1 - x})$.

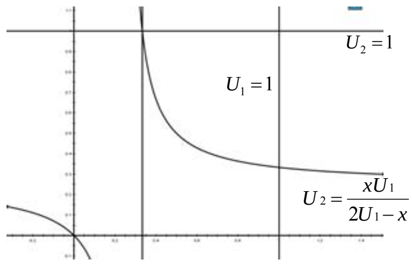


Figure 3.

When $0 < x < 1$, we establish the plane rectangular coordinate system with abscissa axis U_1 and vertical axis U_2 (see Fig. 3.). Note that (U_1, U_2) locate in the field surrounded by $U_1 = 1$, $U_2 = 1$ and abscissa axis, vertical axis (except the origin) is an equally likely event. Then it satisfies the condition of geometric probability model. Since $U_2 \le \frac{xU_1}{2U_1 - x}$ is the field surrounded by $U_2 = \frac{xU_1}{2U_1 - x}$, $U_1 = 1$, $U_2 = 1$ and abscissa axis, vertical axis, then $U_2 = \frac{xU_1}{2U_1 - x}$ intersects with $U_2 = 1$ at the intersection point $(\frac{x}{2-x}, 1)$. Hence

$$P(U_2 \le \frac{xU_1}{2U_1 - x}) = \int_{\frac{x}{2-x}}^{1} \frac{xU_1}{2U_1 - x} dU_1 + \frac{x}{2-x}$$

$$= [\frac{x}{2}U_1 + \frac{x^2}{4}\ln(4U_1 - 2x)]\Big|_{\frac{x}{2-x}}^{1} + \frac{x}{2-x} = x + \frac{x^2}{2}\ln(\frac{2}{x} - 1).$$

Since $0 < H_2 \le 1$,

$$F_{H_2}(x) = \begin{cases} 0 \quad (x \le 0); \\ x + \frac{x^2}{2}\ln(\frac{2}{x} - 1) \quad (0 < x < 1); \\ 1 \quad (x \ge 1); \end{cases}$$

then $$F_{H_2}(E(U)) = F_{H_2}(\frac{1}{2}) = \frac{1}{2} + \frac{\ln 3}{8} = 0.6373,$$

$$p_{H_2}(x) = F_{H_2}{}'(x) = \begin{cases} 0 \quad (x \le 0 \quad or \quad x \ge 1); \\ 1 + x\ln(\frac{2}{x} - 1) + \frac{x}{x-2} \quad (0 < x < 1). \end{cases}$$

The average error is $\bar{e}(H_2) = E(|x - E(x)|) = \int_0^1 \left|x - \frac{1}{2}\right| p_{H_2}(x)dx$.

Let $f(x) = (\frac{x^3}{3} - \frac{x^2}{4})\ln(\frac{2}{x} - 1) - \frac{4}{3}\ln(2 - x) + \frac{2x^2}{3} + \frac{x}{6}$. It can be checked that

$$f'(x) = (x - \frac{1}{2}) p_{H2}(x) \quad (0 < x < 1).$$

Hence $\overline{e}(H_2) = f(x)|_{\frac{1}{2}}^{1} + [-f(x)]|_{0}^{\frac{1}{2}} = 4\ln 2 - \frac{21}{8}\ln 3 + \frac{1}{3} = 0.2221$.

2.2.4. Conclusions

In conclusion, $\overline{e}(A_2) < \overline{e}(G_2) < \overline{e}(H_2) < \overline{e_U}$. It can be seen that choosing different means, the effect of reducing the error may be different.

2.3. $f(x)$-average error for small error numbers

The small error number is defined to be a number which is largely greater than its error. If a set of small error numbers satisfies uniform distribution in the error range, without loss of generality, suppose $X_i = a + bU = b(\frac{a}{b} + U)$. Let $n = \frac{a}{b}$. Then $X_i = b(n + U)$. It is easy to see that *n* is very large for small error numbers. Now we discuss the properties of the means when $n \to +\infty$.

Lemma 1. If the inversion function of $f(x)$ exists, $f(x)$ and $f^{-1}(x)$ are both derivable, then $f^{-1'}[f(x)] \cdot f'(x) = 1$.

Proof. See Fig. 4. $f'(x)$ is equal to the slope k_1 of the tangent line l of $f(x)$ at the point $A(x, f(x))$, and $f^{-1'}[f(x)]$ is equal to the slope k_2 of the tangent line m of $f^{-1}(x)$ at the point $B(f(x), x)$.

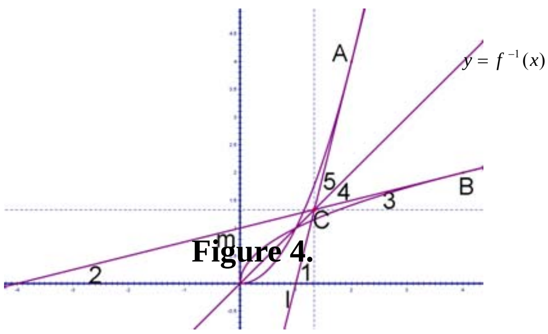


Figure 4.

The images of $y = f(x)$ and $y = f^{-1}(x)$ (resp. A and B) are symmetric with regard to the line $y = x$. Then by the symmetry,

I. If $l // m$, then l、m are both parallel to the line $y = x$. Hence

$$k_1 = k_2 = 1, \text{ and then } k_1 k_2 = 1.$$

II. If l intersects with m, then l, m intersect with the line $y = x$ at a point C. Draw two lines at the point C which are parallel to *x*-axis and *y*-axis, respectively.

Since $\angle 1 = \angle 3 + \angle 4 + \angle 5$, $\angle 2 = \angle 3$, $\angle 4 = \angle 5$, then $\angle 1+\angle 2=2(\angle 3+\angle 4)=2\times 45°=90°$,

and $k_1 \cdot k_2=\tan\angle 1\cdot\tan\angle 2=1$.

Consequently, $f^{-1'}[f(x)]\cdot f'(x)=1$. ■

Lemma 2. Suppose the inversion function of $f(x)$ exists, $f(x)$ and $f^{-1}(x)$ are both derivable. If $f(x)$ is an increasing (resp. decreasing) function over I, then $f^{-1}(x)$ is an increasing (resp. decreasing) function over $M=\{f(x)|x\in I\}$.

Proof. Suppose $f(x)$ is an increasing (resp. decreasing) function over I, $f^{-1}(x)$ is an increasing (resp. decreasing) function over N ($N\subseteq M$). Let $x_0\in N$. Then $f'(x_0)$ and $f^{-1'}[f(x_0)]$ have different signs. Then

$$f'(x_0)\cdot f^{-1'}[f(x_0)]<0,$$

which is a contradiction to Lemma 1. This completes the proof. ■

Lemma 3. If the inverse function of $f(x)$ exists, $f(x)$ and $f^{-1}(x)$ are both derivable, then $\overline{x_f}=f^{-1}[\frac{1}{n}\sum_{i=2}^{n}f(x_i)+\frac{f(x_1)}{n}]$ is monotonically increasing, where x_1 is a variable, and x_2, x_3, . . . , x_n are constant values.

Proof. Since $f(x)$ is a continuous function and has the inverse function, then $f(x)$ is monotonical.

I. $f(x)$ is monotonically increasing.

Then $f^{-1}(x)$ and $\frac{1}{n}\sum_{i=2}^{n}f(x_i)+\frac{f(x_1)}{n}$ are monotonically increasing.

Therefore $f^{-1}[\frac{1}{n}\sum_{i=2}^{n}f(x_i)+\frac{f(x_1)}{n}]$ is monotonically increasing, i.e., $\overline{x_f}$ is monotonically increasing.

II. $f(x)$ is monotonically decreasing.

Then $f^{-1}(x)$ and $\frac{1}{n}\sum_{i=2}^{n} f(x_i)+\frac{f(x_1)}{n}$ are monotonically decreasing.

Therefore $f^{-1}[\frac{1}{n}\sum_{i=2}^{n} f(x_i)+\frac{f(x_1)}{n}]$ is monotonically increasing, i.e., $\overline{x_f}$ is monotonically increasing.

The result follows. ■

Theorem 8. Suppose a function $f(x)$ has the inverse function $f^{-1}(x)$. If $f(x)$ and $f^{-1}(x)$ are both derivable, then $L=\lim_{n\to+\infty}(\overline{x_f}-n)$ and $\frac{U_1+U_2}{2}$ have the same distribution, where $x_1=n+U_1$, $x_2=n+U_2$, and $\overline{x_f}$ is the $f(x)$-average mean of x_1 and x_2.

Proof. Note that $f^{-1}[\frac{f(n)+f(n)}{2}]<\overline{x_f}\le f^{-1}[\frac{f(n+1)+f(n+1)}{2}]$, i.e., $n<\overline{x_f}\le n+1$.

Then
$$0<\lim_{n\to+\infty}(\overline{x_f}-n)\le 1,$$
and
$$F_L(x)=0\ (x\le 0),\quad F_L(x)=1\ (x\ge 1).$$

For $0<x<1$, we have
$$F_L(x)=P(f^{-1}[\frac{f(x_1)+f(x_2)}{2}]-n\le x)$$
$$=P(x_2\le f^{-1}[2f(n+x)-f(x_1)])\ \text{(By Lemma 3)}.$$

Let two dimensional Cartesian coordinate system consist of the abscissa axis x_1 and vertical axis x_2. Note that (x_1,x_2) locate in the field surrounded by $x=n$, $y=n$, $x=n+1$ and $y=n+1$ (except (n,n)) is an equally likely event. Then the conditions of geometric probability model are satisfied.

$x_2\le f^{-1}[2f(n+x)-f(x_1)]$ is the region surrounded by $x=n$, $y=n$, $x=n+1$, $y=n+1$ and the below area of $f^{-1}[2f(n+x)-f(x_1)]$.

Then $\{f^{-1}[2f(n+x)-f(x_1)]\}_{x_1}'=-f^{-1'}[2f(n+x)-f(x_1)]\cdot f'(x_1)$.

When $n\to+\infty$, then $f(n+x)=f(n)$, $f(x_1)=f(n)$, $f'(x_1)=f'(n)$ and
$$\{f^{-1}[2f(n+x)-f(x_1)]\}'=-f^{-1'}[f(n)]\cdot f'(n)=-1.$$

Hence $f^{-1}[2f(n+x)-f(x_1)]$ is the straight line l :

$x_2-(n+x)=(n+x)-x_1$, where the line l passes the point $(n+x,n+x)$ with slope -1.

Then the line passes point $A(n,n+2x)$ and point $E(n+1,n+2x-1)$.

When $0<x\le\frac{1}{2}$ (see Figure 5.(a)), we obtain

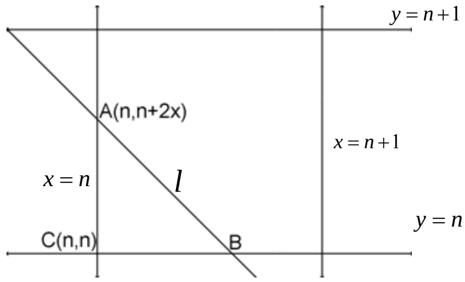


Figure 5.(a)

$P(x_2\le f^{-1}[2f(n+x)-f(x_1)])=S_{\Delta ABC}/1=\frac{1}{2}(2x)^2=2x^2$.

When $\frac{1}{2}<x<1$ (see Figure 5.(b)), we have

$$P(x_2\le f^{-1}[2f(n+x)-f(x_1)])=(1-S_{\Delta DEF})/1=1-\frac{(2-2x)^2}{2}=1-2(1-x)^2,$$

i.e.,

$$F_L(x)=\begin{cases}0 & (x\le 0);\\ 2x^2 & (0<x\le\frac{1}{2});\\ 1-2(1-x)^2 & (\frac{1}{2}<x<1);\\ 1 & (x\ge 1).\end{cases}$$

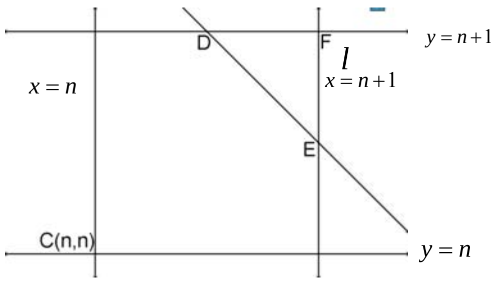


Figure 5.(b)

Hence $F_L(x)=F_{A_2}(x)$. ■

It is shown that the distribution of $\lim_{n\to+\infty}(\overline{x_f}-n)$ is independent from $f(x)$. Combining Theorem 5, for any continuous function $f(x)$ which has the inverse function and $x_1=n+U_1$, $x_2=n+U_2$, we have

$$\lim_{n\to+\infty}\bar{e}(\overline{x_f})=\bar{e}(A_2)=\frac{1}{6}.$$

Example 5. Let R_1 and R_2 be two same resistances labeled by "$2k\Omega$ ($\pm 0.1\%$)". Suppose R_1 and R_2 are both uniform distribution within the error scope. Try to estimate the error of the resistance R raised from the parallel of R_1 and R_2.

Solution. Since $R_1,R_2\in(1998,2002]$ (Unit: Ω), then

$$R_1 = 1998 + 4U_1 = 4(499.5 + U_1) .$$

Similarly, we obtain $R_2 = 1998 + 4U_2 = 4(499.5 + U_2)$.

Let $f(x) = \frac{1}{x}$ and $\overline{R_f}$ be the $f(x)$-mean of R_1 and R_2. Then

$$R = \frac{R_1 R_2}{R_1 + R_2} = \frac{\overline{R_f}}{2} .$$

Note that $n = 499.5$ is relatively large. Therefore $\frac{R}{2} - 499.5 \approx \frac{U_1 + U_2}{2}$, i.e., $R \approx 2A_2 + 999$. Hence $\overline{e_R} \approx 2\overline{e_{A_2}} = 2 \times \frac{1}{6} = \frac{1}{3}\Omega$.

3. The effect of Large numbers and Small numbers on means

For four numbers 1, 1, 1, 4, the x-mean is 1.75 and the $\ln x$-mean is $\sqrt{2} = 1.414$. Obviously, the effects of the two means raised by extreme value（4）are different. Extreme value（4） has more affect on the x-mean than the $\ln x$-mean. Naturally, a problem is raised that how to compare the effect of a kind of mean from extreme value conditions?

Theorem 9. Let $f(x)$ be a continuous function and $\overline{U_f}$ the $f(x)$-mean of U_1 and U_2. Then $x^2 \le F_{\overline{U_f}}(x) \le 1 - (1-x)^2$.

Proof. When $U_1 \le x$ and $U_2 \le x$, then

$$\overline{x_f} \le f^{-1}[\frac{f(x) + f(U_2)}{2}] \le f^{-1}[\frac{f(x) + f(x)}{2}] = x \text{ and}$$

$$P(\overline{U_f} \le x) \ge P(U_1 \le x) \cdot P(U_2 \le x) = x^2 .$$

When $U_1 > x$ and $U_2 > x$, then $\overline{U_f} > f^{-1}[\frac{f(x) + f(U_2)}{2}] > f^{-1}[\frac{f(x) + f(x)}{2}] = x$ and

$$P(\overline{U_f} > x) = 1 - P(\overline{U_f} \le x) \ge P(U_1 > x) \cdot P(U_2 > x) = (1-x)^2 ,$$

i.e., $P(\overline{U_f} \le x) \le 1 - (1-x)^2$.

Hence $x^2 \le F_{\overline{U_f}}(x) \le 1 - (1-x)^2$. ■

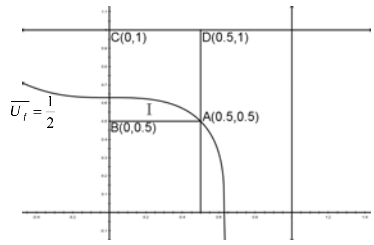
Especially, when $x = E(U)$, we have $\frac{1}{4} \le P(\overline{U_f} \le x) \le \frac{3}{4}$.

Let $P(\overline{U_f} \le E(U)) = \frac{1}{4}$ and if one of U_1 and U_2 is larger than $E(U)$, then

$\overline{U_f}$ is larger than $E(U)$. It is shown that $\overline{U_f}$ has enormous implications by large numbers. Let $P(\overline{U_f} \le E(U)) = \frac{3}{4}$ and if one of U_1 and U_2 is smaller than $E(U)$, then $\overline{U_f}$ is smaller than $E(U)$. It is shown that $\overline{U_f}$ has enormous implications by small numbers.

Let two dimensional Cartesian coordinate system consist of the abscissa axis U_1 and vertical axis U_2. (See Figure 6.). Note that (U_1, U_2) locate in the field surrounded by $U_1 = 1$ 、 $U_2 = 1$ and the two axis (except the origin) is an equally likely event. Then the conditions of geometric probability model are satisfied. Generally，let $d = P(\overline{U_f} \le E(U))$, $U_1 \le E(U)$ and $U_2 \ge E(U)$, then the probability of $\overline{U_f} \le E(U)$ is equal to

$$p = \frac{S_{\mathrm{I}}}{S_{ABCD}} = \frac{\frac{1}{2}(d - \frac{1}{4})}{\frac{1}{4}} = 2d - \frac{1}{2}$$



and the probability of $\overline{U_f} > E(U)$ is $1 - p$.

Figure 6.

Definition 3. The impact coefficient of the $f(x)$ -mean is defined as $p_f = 2F_{\overline{U_f}}(\frac{1}{2}) - \frac{1}{2}$ （$\overline{U_f} = f^{-1}[\frac{f(U_1) + f(U_2)}{2}]$） .

It is obvious that $0 \le p \le 1$. If p_f is more larger, then the $f(x)$-mean acted by small numbers has more impactions; if p_f is more smaller, then the $f(x)$-mean acted by larger numbers has more impactions. If $p_f = 0.5$, the $f(x)$-mean acted by larger numbers and small numbers has the same impactions.

4. The necessary and sufficient condition of comparing values and identical equality of means

4.1. Comparing values of means

Mean inequality shows the relationship of sizes between arithmetic mean, geometric mean and harmonic mean. For the generalized $f(x)$ -mean and

$g(x)$-mean, what are the relationship of their sizes?

Theorem 10. If $f(x)$ is monotonically increasing and $f[g^{-1}(x)]$ is a concave function, then $\overline{x_f} \le \overline{x_g}$ with equality holds if and only if $x_1 = x_2 = \cdots = x_n$.

Proof. Let $h(x) = f[g^{-1}(x)]$.

Since $g(x)$ is monotonically increasing, then $g^{-1}(x)$ is monotonically increasing.

Then $h(x)$ and $h^{-1}(x)$ are both monotonically increasing.

Note that $h(x)$ is a concave function. By Jensen Inequality,

$h(\frac{1}{n}\sum_{i=1}^{n} y_i) \ge \frac{1}{n}\sum_{i=1}^{n} h(y_i)$, where the equality holds if and only if $y_1 = y_2 = \cdots = y_n$.

Then $$\frac{1}{n}\sum_{i=1}^{n} y_i \ge h^{-1}[\frac{1}{n}\sum_{i=1}^{n} h(y_i)].$$

Suppose $x_i = g^{-1}(y_i)$ ($i = 1,2,...,n$) , then $y_i = g(x_i)$.

We have $\frac{1}{n}\sum_{i=1}^{n} g(x_i) \ge h^{-1}[\frac{1}{n}\sum_{i=1}^{n} f(x_i)]$.

Then $f\{g^{-1}[\frac{1}{n}\sum_{i=1}^{n} g(x_i)] \ge \frac{1}{n}\sum_{i=1}^{n} f(x_i)$, and $g^{-1}[\frac{1}{n}\sum_{i=1}^{n} g(x_i)] \ge f^{-1}[\frac{1}{n}\sum_{i=1}^{n} f(x_i)]$, i.e., $\overline{x_f} \le \overline{x_g}$ with equality holds if and only if $x_1 = x_2 = \cdots = x_n$. ■

If $g(x)$ is monotonically decreasing, then the same conclusion is also obtained.

Similarly, it is easy to obtain the following:

If $f(x)$ is monotonically decreasing and $f[g^{-1}(x)]$ is a concave function, then

$$\overline{x_f} \ge \overline{x_g}.$$

If $f(x)$ is monotonically increasing and $f[g^{-1}(x)]$ is a convex function, then

$$\overline{x_f} \ge \overline{x_g}.$$

If $f(x)$ is monotonically decreasing and $f[g^{-1}(x)]$ is a convex function, then

$$\overline{x_f} \le \overline{x_g}.$$

(The equality holds if and only if $x_1 = x_2 = \cdots = x_n$.)

Example 5. (Mean inequality) $\frac{1}{n}\sum_{i=1}^{n} x_i \ge (\prod_{i=1}^{n} x_i)^{\frac{1}{n}} \ge \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$ ($x_i > 0$) .

Proof. Let $a(x) = x$, $g(x) = \ln x$ and $h(x) = \frac{1}{x}$. By Theorem 1, it only needs to prove $\overline{x_a} \ge \overline{x_g} \ge \overline{x_h}$.

Since $a''[g^{-1}(x)] = (e^x)'' = e^x > 0$, then $a[g^{-1}(x)]$ is a convex function.

Note that $a(x)$ is monotonically increasing. Then we have $\overline{x_a} \ge \overline{x_g}$.

By $g''[h^{-1}(x)] = (\ln\frac{1}{x})'' = (-\frac{1}{x})' = \frac{1}{x^2} > 0$, $g[h^{-1}(x)]$ is a convex function.

And $g(x)$ is monotonically increasing. Then $\overline{x_g} \ge \overline{x_h}$ holds.

From the above discussions, we have $\overline{x_a} \ge \overline{x_g} \ge \overline{x_h}$.

The proof is completed. ■

4.2. The relationship of sizes and impact coefficients of means

Theorem 11. If $\overline{x_f} \le \overline{x_g}$ holds, then $p_f \ge p_g$.

Proof. If U_1 and U_2 satisfy $\overline{U_g} \le \frac{1}{2}$, then U_1 and U_2 shall satisfy $\overline{U_f} \le \frac{1}{2}$.

Then $F_{\overline{U_f}}(\frac{1}{2}) \ge F_{\overline{U_g}}(\frac{1}{2})$.

Hence $p_f \ge p_g$. ■

By Theorem 8 and its generalized version and Theorem 9, we obtain the following conclusion:

Corollary 1. Let $h(x) = f[g^{-1}(x)]$. If $f'(x)\cdot h''(x) < 0$, then $p_f \ge p_g$; if $f'(x)\cdot h''(x) > 0$, then $p_f \le p_g$.

Example 6. Let $f(x) = x^n$ ($x > 0$, $n > 0$). Then p_f is monotonically decreasing on n.

Proof. Let $n_2 > n_1 > 0$, $f_1(x) = x^{n_1}$ and $f_2(x) = x^{n_2}$. Then

$$h(x) = f_2[f_1^{-1}(x)] = x^{\frac{n_2}{n_1}}.$$

Therefore, $h^{''}(x) = \frac{n_2}{n_1}(\frac{n_2}{n_1} - 1)x^{\frac{n_2}{n_1}-2}$.

Since $\frac{n_2}{n_1} > 1$ and $x > 0$, then $h^{''}(x) > 0$ holds.

Note that $f_2^{'}(x) = n_2 x^{n_2 - 1} > 0$. Then $f^{'}(x) \cdot h^{''}(x) > 0$.

We have $p_{f_2} \le p_{f_1}$.

Hence p_f is monotonically decreasing on n. ■

For $f(x) = x^n$ （$n > 1$） and $p_f = p$ (p is a constant), the value of n may be searched by the following program（Pas Language）:

```
{$N+}
var i,j,d:longint;
      n,a,b,p,dx,s:extended;
function f(x,n:extended):extended;
begin
   if x=0 then f:=exp(ln(0.5)*(n-1)/n)
   else f:=exp(ln(exp(ln(0.5)*(n-1))-exp(ln(x)*n))/n);
end;
begin
   readln(p,a,b,d);
   repeat
      n:=(a+b)/2;
      dx:=f(0,n)/d;
      s:=0;
      for j:=1 to d-1 do
      begin
         s:=s+f(j*dx,n)*dx;
      end;
      s:=2*s-0.5;
      if s=p then
      begin
         writeln(n);
         break;
      end;
      if s>p then a:=(a+b)/2
```

```
    else b:=(a+b)/2
  until b-a<1E-5;
  writeln(n:0:4);
end.
```

In the practical application, we may reference the following table. By selecting the functions which have small impact coefficients, the impact of extreme small value can be reduced; By selecting the functions which have large impact coefficients, the impact of extreme large value can be reduced.

$f(x)$ $(x>0)$	$x^{6.46}$	$x^{3.02}$	$x^{1.88}$	$x^{1.33}$	x	$\ln x$	$\frac{1}{x}$
p_f	0.1	0.2	0.3	0.4	0.5	0.693	0.775

4.3. The necessary and sufficient condition of $\overline{x_f} \equiv \overline{x_g}$

Lemma 4. ([3]) If $f''(x)=0$, then $f(x)$ is constant or a linear function.

Theorem 12. If $f(x)$ and $g(x)$ are both continuous functions and $\overline{x_f} \equiv \overline{x_g}$, then

$$g(x)=af(x)+b \quad (a \neq 0).$$

Proof. Let $h(x)=f[g^{-1}(x)]$. Suppose $h''(x)\neq 0$, then $h(x)$ is a concave or convex function. Since there exists $\overline{x_f}$, $f(x)$ has the inverse function.

Note that $f(x)$ is a continuous function. Then $f(x)$ is a monotonical function.

By Theorem 8 and its generalized version,

for $x_1 = x_2 = \cdots = x_{n-1} \neq x_n$, we have $\overline{x_f} \neq \overline{x_g}$, a contradiction. Hence $h''(x)=0$.

By Lemma 4, we may assume

$$h(x)=\frac{x-b}{a} \quad (a\neq 0) \text{ or } h(x)=c \quad (c \text{ is a constant}).$$

Suppose $h(x)=c$, i.e., $f[g^{-1}(x)]=c$. Then $g^{-1}(x)=f^{-1}(c)$ and $g(x)=f(c)$.

Hence $g(x)$ has no inverse functions. Therefore, $\overline{x_g}$ does not exist, a contradiction.

Then $f[g^{-1}(x)]=\frac{x-b}{a}$.

We have $g^{-1}(x) = f^{-1}(\frac{x-b}{a}) = y$ and $x = af(y)+b$.

Hence $g(x) = af(x)+b$ （$a \neq 0$）. ■

Combining Theorem 2, we obtain that for the continuous functions $f(x)$ and $g(x)$, if $f(x)$ and $g(x)$ have both the inverse function, then $\overline{x_f} \equiv \overline{x_g}$ holds if and only if $g(x) = af(x)+b$ （$a \neq 0$）.

Corollary 2. Let $h(x) = f[g^{-1}(x)]$. Then $f^{'}(x)\cdot h^{''}(x) < 0$ holds if and only if $\overline{x_f} \le \overline{x_g}$ (The equality holds if and if $x_1 = x_2 = \cdots = x_n$);

And $f^{'}(x)\cdot h^{''}(x) > 0$ holds if and only if $\overline{x_f} \ge \overline{x_g}$ (The equality holds if and only if $x_1 = x_2 = \cdots = x_n$).

Proof. By Theorem 8 and its generalized version, let $h(x) = f[g^{-1}(x)]$. If $f^{'}(x)\cdot h^{''}(x) < 0$ holds, then $\overline{x_f} \le \overline{x_g}$; if $f^{'}(x)\cdot h^{''}(x) > 0$ holds, then $\overline{x_f} \ge \overline{x_g}$. (The equalities hold if and only if $x_1 = x_2 = \cdots = x_n$.)

If $\overline{x_f} \le \overline{x_g}$ holds（The equality holds if and only if $x_1 = x_2 = \cdots = x_n$）, then we can suppose $f^{'}(x)\cdot h^{''}(x) \ge 0$.

By the proof course of Theorem 10, we have $h^{''}(x) \neq 0$. Note that $f^{'}(x) \neq 0$（Since the definition of $\overline{x_f}$ requires that $f(x)$ has the inverse function, $f(x)$ is not a constant.）. Then $f^{'}(x)\cdot h^{''}(x) \neq 0$， i.e., $f^{'}(x)\cdot h^{''}(x) > 0$.

From the above discussions, $\overline{x_f} \ge \overline{x_g}$ holds（The equality holds if and only if $x_1 = x_2 = \cdots = x_n$.）, a contradiction. Hence

$$f^{'}(x)\cdot h^{''}(x) < 0.$$

Therefore, we have

$f^{'}(x)\cdot h^{''}(x) < 0$ holds if and only if $\overline{x_f} \le \overline{x_g}$（The equality holds if and only if $x_1 = x_2 = \cdots = x_n$.）.

Similarly, $f^{'}(x)\cdot h^{''}(x)>0$ holds if and only if $\overline{x_f}\geq\overline{x_g}$ （The equality holds if and only if $x_1=x_2=\cdots=x_n$.）. ■

References

[1] S.S. Mao, Y.M. Cheng, X.L. Pu, *An Introduction on Probability and Statistics*, Higher Education Press, 2004.

[2] X.G. Huang, *The convex functions and Jensen Inequality*, ShangHai Education Publishing House, 1991.

[3] Z.Q Ye, *Higher Mathematics (First volume)*, Higher Education Press, 2004.