УДК 519.213 doi: 10.26906/SUNZ.2018.2.039

В.Ю. Дубницкий, О.Е. Петренко, А.И. Ходырев

Харьковский учебно-научный институт ГВУЗ «Университет банковского дела», Харьков

ВЫЧИСЛЕНИЕ ЭНТРОПИИ ФУНКЦИИ КВАНТИЛЕЙ (ЭНТРОПИИ ВАСИЧЕКА)

Поставлена задача определения численного значения энтропии функции квантилей (энтропии Васичека) для непрерывной случайной величины. Для показательного распределения, распределения Чампернауна, логистического распределения, распределения арксинуса, распределения Рэлея, распределения Парето, распределения Лапласа получены в явном виде выражения для определения численных значений энтропии функции квантилей (энтропии Васичека). Для распределения Коши, распределения минимального значения, распределения максимального значения, двойного показательного распределения и распределения Вейбулла описана процедура решения поставленной задачи численными методами. Предложена единица измерения энтропии Васичека - вит.

Ключевые слова: *Васичек, энтропия, показательное распределение, распределение Чампернауна, логистическое распределение, распределение арксинуса, распределение Рэлея, распределение Парето, распределение Лапласа, распределение Коши, распределение минимального значения, распределение максимального значения, двойное показательное распределение, распределение Вейбулла, численные методы.*

Введение

Пусть для непрерывной случайной величины X известны функция распределения $F(x)$ и плотность распределения $f(x)$. Энтропией Шеннона этой случайной величины называют функционал вида:

$$H_{sh} = -\int_{-\infty}^{\infty} \ln f(x) f(x) dx . \qquad (1)$$

Формальные свойства этого функционала подробно описаны в работе [1]. Его применение в технике связи изложено в работах [2, 3, 4].

Для дальнейшего изложения используем понятия квантиля. В работе [5] квантиль x_p уровня p, $0 < p < 1$, определён как корень уравнения

$$F\left(x_p\right) = p, \qquad (2)$$

откуда следует, что:

$$x_p = F^{-1}(p) . \qquad (3)$$

В работе [6] было введено условие вида:

$$H_v = \int_0^1 \ln\left(\frac{d}{dx} F^{-1}(p) dp\right). \qquad (4)$$

Это условие получило название энтропии функции квантилей или энтропии Васичека. Вид энтропии Шеннона для большинства наиболее распространённых видов законов распределения дискретной и непрерывной случайной величины X приведен в работе [7].

Так как в доступной авторам данного сообщения литературе отсутствует наименование единицы измерения энтропии Васичека, то, по аналогии с единицей измерения энтропии Шеннона, предлагаем называть её «вит».

Анализ литературы. В работе [5] было отмечено, что функция квантилей тесно связана с порядковыми статистиками. Поэтому условие (4) было использовано при построении различных критериев соответствия эмпирических функций распределения их теоретическим аналогам. Обоснование это подхода дано в работе [6]. На алгоритмическом уровне применение этого критерия описано в работе [7]. Результаты его применения, полученные в последние годы, изложены в работах [8, 9].

Для наиболее распространённых функций распределения дискретных и непрерывных случайных величин в работе [10] приведены выражения для энтропии Шеннона. Для энтропии Васичека аналогичных исследований авторам данного сообщения найти не удалось.

Постановка задачи: вычисление энтропии квантилей (энтропии Васичека) для конкретных законов распределения непрерывной случайной величины X.

Полученные результаты

В настоящей работе энтропия Васичека определена для тех случаев, когда распределение непрерывной случайной величины X соответствует: 1) показательному закону, 2) распределению Чампернауна, 3) логистическому закону, 4) распределению арксинуса, 5) распределению Рэлея, 6) распределению Парето, 7) распределению Лапласа, 8) распределению Коши, 9) распределению минимального значения, 10) распределению максимального значения, 11) двойному показательному распределению, 12) распределению Вейбулла. Необходимые сведения о функциях распределения и функциях квантилей, свойственных этим законам, приведены в справочнике [11] и использованы авторами данного сообщения. Для случаев 1) …7) решение поставленной задачи получено в явном виде. Для случаев 8) …12) решение поставленной задачи получено численными методами. При решении задач использована система компьютерной алгебры DERIVE 6 [12].

Для законов распределения, допускающих получение выражения энтропии Васичека в явном виде, получены следующие результаты.

Рассмотрим решение поставленной задачи для случайной величины X, имеющей показательное распределение. Функция распределения этого закона имеет вид:

$$F(x) = 1 - \exp(-\lambda x),\ x > 0. \tag{5}$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x) = -\frac{1}{\lambda}\ln(1-p). \tag{6}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}\left(-\frac{1}{\lambda}\ln(1-p)\right) = \frac{1}{\lambda(1-p)}. \tag{7}$$

Таким образом, энтропия Васичека для показательного распределения будет равна величине:

$$H_V(p) = \int_0^1 \ln\left[\frac{1}{\lambda(1-p)}\right]dp = \ln\left(\frac{1}{\lambda+1}\right), \text{(вит)}. \tag{8}$$

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение Чампернауна. Функция распределения этого закона имеет вид:

$$F(X) = \frac{2}{\pi}\operatorname{arctg}\left[\exp\left(\alpha\cdot(x-\mu)\right)\right],\ -\infty < x < \infty. \tag{9}$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x) = \mu + \frac{1}{\alpha}\ln\left(\operatorname{tg}\frac{\pi p}{2}\right). \tag{10}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}\left(\mu + \frac{1}{\alpha}\ln\left(\operatorname{tg}\frac{p\pi}{2}\right)\right) = \frac{\pi}{\alpha\sin(p\pi)}. \tag{11}$$

Таким образом, энтропия Васичека для распределения Чампернауна будет равна величине:

$$H_V(p) = \int_0^1 \ln\left[\frac{\pi}{\alpha\sin(p\pi)}\right]dp = \ln\left(\frac{2\pi}{\alpha}\right), \text{(вит)}. \tag{12}$$

Рассмотрим решение поставленной задачи для случайной величины X, имеющей логистическое распределение. Функция распределения этого закона имеет вид:

$$F(x) = \frac{1}{1+\exp\left(-\dfrac{x-\mu}{\lambda}\right)} = \frac{1}{2}\left[1+\operatorname{th}\left(\frac{x-\mu}{\lambda}\right)\right],\ -\infty < x < \infty. \tag{13}$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x) = \mu - \lambda\ln\left((1-p)/p\right). \tag{14}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}\left(\mu - \lambda\ln\frac{1-p}{p}\right) = \frac{\lambda}{p(1-p)}. \tag{15}$$

Таким образом, энтропия Васичека для логистического распределения будет равна величине:

$$H_V(p) = \int_0^1 \ln\left[\frac{\lambda}{p(1-p)}\right]dp = \ln\lambda + 2, \text{(вит)}. \tag{16}$$

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение арксинуса. Функция распределения этого закона имеет вид:

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\arcsin\frac{x-\mu}{\lambda},\quad \frac{2\mu-\pi\lambda}{2} \le x\,\frac{2\mu+\pi\lambda}{2}. \tag{17}$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x) = \mu + \lambda\sin\left[\pi\left(p-1/2\right)\right]. \tag{18}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}\left(\mu + \lambda\sin\left[\pi\left(p-\frac{1}{2}\right)\right]\right) = \pi\lambda\sin(p\pi). \tag{19}$$

Таким образом, энтропия Васичека для распределения арксинуса будет равна величине:

$$H_V(p) = \int_0^1 \ln\left[\pi\lambda\sin(p\pi)\right]dp = \ln\left(\frac{\pi\lambda}{2}\right), \text{(вит)}. \tag{20}$$

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение Рэлея. Функция распределения этого закона имеет вид:

$$F(x) = 1 - \exp\left[-x^2/\left(2a^2\right)\right],\ x > 0. \tag{21}$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x) = a\sqrt{-2\ln(1-p)}. \tag{22}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}a\sqrt{-2\ln(1-p)} = \frac{a\sqrt{2}}{2(1-p)\sqrt{-\ln(1-p)}}. \tag{23}$$

Таким образом, энтропия Васичека для распределения арксинуса будет равна величине:

$$H_V(p) = \int_0^1 \ln\left[\frac{a\sqrt{2}}{2(1-p)\sqrt{-\ln(1-p)}}\right]dp =$$

$$= \ln\left[\frac{1}{(1-p)\sqrt{-\ln(1-p)}}\right] - \frac{\ln 2}{2} - 1,\quad \text{(вит)}. \tag{24}$$

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение Парето. Функция распределения этого закона имеет вид:

$$F(x) = 1 - \left(x_0/x\right)^{\alpha}, \; x > x_0 . \tag{25}$$

Функция квантилей для этого распределения имеет вид

$$F^{-1}(x) = x_0 \big/ (1-p)^{1/\alpha} . \tag{26}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}\frac{x_0}{(1-p)^{1/\alpha}} = \frac{x_0(1-p)^{-(\alpha+1)/\alpha}}{\alpha} . \tag{27}$$

Таким образом, энтропия Васичека для распределения Парето будет равна величине:

$$\begin{aligned} H_v(p) &= \int_0^1 \ln\left(x_0(1-p)^{-(\alpha+1)/\alpha} \big/ \alpha\right) dp = \\ &= \ln\left(x_0/\alpha\right) + (\alpha+1)/\alpha, \quad (\text{вит}). \end{aligned} \tag{28}$$

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение Лапласа. Функция распределения этого закона имеет вид:

$$F(x) = \begin{cases} 0.5\exp\left(\lambda(x-\mu)\right), & x \le \mu; \\ 1 - 0.5\exp\left(\lambda(x-\mu)\right), & x \ge \mu; \end{cases} \tag{29}$$

при условии, что $-\infty < x < \infty$.

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x) = \begin{cases} \mu + \dfrac{\ln(2p)}{\lambda}, & 0 < p < 0,5; \\ \mu - \dfrac{\ln\left(2(1-p)\right)}{\lambda}, & 0.5 < p < 1. \end{cases} \tag{30}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}F^{-1}(x) = \begin{cases} 1/(p\lambda), & 0 < p < 0,5; \\ 1/\left(\lambda(1-p)\right), & 0,5 < p < 1. \end{cases} \tag{31}$$

Таким образом, энтропия Васичека для распределения Лапласа будет равна величине:

$$\begin{aligned} H_v(p) &= \int_0^{0,5} \ln\left(\frac{1}{p\lambda}\right) dp + \int_{0,5}^1 \ln\frac{1}{\lambda(1-p)} dp = \\ &= \ln(2/\lambda) + 1 \quad (\text{вит}). \end{aligned} \tag{32}$$

Далее изложим результаты определения энтропии Васичека, полученные численным интегрированием.

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение Коши. Функция распределения этого закона имеет вид:

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}\frac{x-\mu}{\lambda}, \; -\infty < x < \infty . \tag{33}$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x) = \mu + \lambda \operatorname{tg}\frac{\pi(2p-1)}{2} . \tag{34}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}F^{-1}(x) = \frac{2\pi\lambda}{1-\cos(2p\pi)} . \tag{35}$$

Таким образом, энтропия Васичека для распределения Коши будет равна величине интеграла вида:

$$\int_0^1 \ln\frac{2\pi\lambda}{1-\cos(2p\pi)} dp , \; (\text{вит}). \tag{36}$$

Численные значения энтропии Васичека для распределения Коши при $\mu = 0$ и $\lambda = 0,2(0,2)0,8$ приведены в табл. 1.

Таблица 1

Численные значения энтропии Васичека для распределения Коши

λ	0,2	0,4	0,6	0,8
H_v, (вит).	0,912	1,612	2,018	2,305

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение минимального значения. Функция распределения этого закона имеет вид:

$$F(x) = 1 - \exp\left[-\exp\frac{x-\mu}{\lambda}\right], -\infty < x < \infty . \tag{37}$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(p) = \mu + \lambda \ln\left(-\ln(1-p)\right). \tag{38}$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}F^{-1}(x) = \frac{\lambda}{(p-1)\ln(1-p)} . \tag{39}$$

Таким образом, энтропия Васичека для распределения минимального значения будет равна величине интеграла вида:

$$H_v = \int_0^1 \ln\left[\frac{\lambda}{(p-1)\cdot(1-p)}\right] dp , \; (\text{вит}). \tag{40}$$

Численные значения энтропии Васичека для распределения минимального значения при $\mu = 8$ и $\lambda = 1$, (0,5) 2,5 приведены в табл. 2.

Таблица 2

Численные значения энтропии Васичека для распределения минимального значения

λ	1,0	1,5	2,0	2,5
H_v, (вит).	1,577	1,982	2,270	2,493

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение максимального значения. Функция распределения этого закона имеет вид:

$$F(x)=\exp\left[-\exp\left[-\frac{(x-\mu)}{\lambda}\right]\right],\ -\infty<x<\infty. \quad (41)$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(p)=\mu-\lambda\ln(-\ln p). \quad (42)$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}F^{-1}(x)=-\frac{\lambda}{p\ln p}. \quad (43)$$

Таким образом, энтропия Васичека для распределения максимального значения будет равна величине интеграла вида:

$$H_V(p)=\int_0^1\ln\left(-\frac{\lambda}{p\ln p}\right)dp,\ (\text{вит}). \quad (44)$$

Численные значения энтропии Васичека для распределения максимального значения при $\mu=4$ и $\lambda=1$ (0,5) 2,5 приведены в табл. 3.

Таблица 3

Численные значения энтропии Васичека для распределения максимального значения

λ	1,0	1,5	2,0	2,5
H_V, (вит).	1,577	1,982	2,270	2,493

Заметим, что для этих двух функций распределения энтропия Васичека не зависит от параметра положения μ, но зависит от параметра масштаба $\lambda>0$.

Также следует отметить совпадение численных значений энтропии Васичека для распределений минимального и максимального значений.

Рассмотрим решение поставленной задачи для случайной величины X, имеющей двойное показательное распределение. Функция распределения этого закона имеет вид:

$$F(x)=\exp(-\mu\exp(-\lambda x)),\ -\infty<x<\infty. \quad (45)$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(x)=\frac{\ln\mu-\ln(-\ln p)}{\lambda}. \quad (46)$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}F^{-1}(x)=-\frac{1}{p\lambda\ln p}. \quad (47)$$

Таким образом, энтропия Васичека для двойного показательного распределения будет равна величине интеграла вида:

$$H_V=\int_0^1\ln\left(-\frac{1}{p\lambda\ln p}\right)dp,\ (\text{вит}). \quad (48)$$

Численные значения энтропии Васичека для двойного показательного распределения при $\lambda=0{,}3$ (0,3) 1,2 приведены в табл. 4.

Таблица 4

Численные значения энтропии Васичека для двойного показательного распределения

λ	0,3	0,6	0,9	1,2
H_V, (вит).	2,781	2,088	1,682	1,395

Рассмотрим решение поставленной задачи для случайной величины X, имеющей распределение Вейбулла.

Функция распределения этого закона имеет вид:

$$F(x)=1-\exp\left[-(x/a)^c\right],\ x>0. \quad (49)$$

Функция квантилей для этого распределения имеет вид:

$$F^{-1}(p)=a\left[-\ln(1-p)\right]^{1/c}. \quad (50)$$

Производная функции квантилей по переменной *p* в этом случае примет вид:

$$\frac{d}{dp}F^{-1}(x)=a\left[-\ln(1-p)\right]^{(1-c)/c} \quad (51)$$

Таким образом, энтропия Васичека для распределения Вейбулла будет равна величине интеграла вида:

$$H_V=\int_0^1\ln\left[a\left[-\ln(1-p)\right]^{(1-c)/c}\right]dp,\ (\text{вит}). \quad (52)$$

Изменение величины энтропии Васичека для распределения при Вейбулла при значении его параметров: a=1 (1) 4 и c=1 (0,5) 2,5, показана на рис. 1.

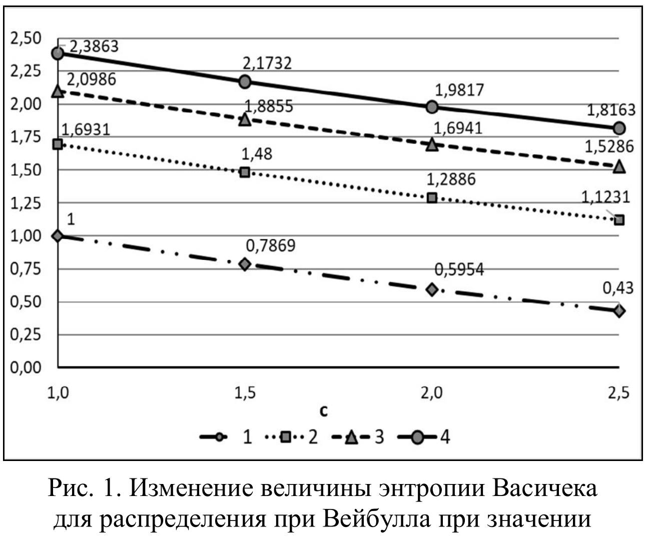

Рис. 1. Изменение величины энтропии Васичека для распределения при Вейбулла при значении его параметров: a=1 (1) 4 и c=1 (0,5) 2,5

В работе не рассмотрено решение поставленной задачи для таких столь распространённых распределений, как нормальное, логарифмически нормальное и гамма-распределение. Это связано с тем, что возникающие при этом вычислительные затруднения, по мнению авторов, должны стать предметом дальнейшего рассмотрения.

Выводы

1. Поставлена задача определения численного значения энтропии квантилей (энтропии Васичека) для непрерывной случайной величины.
2. Для показательного распределения, распределения Чампернауна, логистического распределения, распределения арксинуса, распределения Рэлея, распределения Парето, распределения Лапласа получены в явном виде выражения для определения численных значений энтропии квантилей (энтропии Васичека).
3. Для распределения Коши, распределения минимального значения, распределения максимального значения, двойного показательного распределению и распределения Вейбулла описана процедура решения поставленной задачи численными методами.
4. Предложена единица измерения энтропии Васичека – *вит*.

Список литературы

1. Мартин Н. Математическая теория энтропии. / Н. Мартин, Дж., Ингленд. – Москва: Мир,1988. – 350с.

2. Кузьмин И.В. Основы теории информации и кодирования. / И.В. Кузьмин, В.А., Кедрус. – К.: «Вища школа», 1986. – 238 с.

3. Лифшиц Н.А. Вероятностный анализ систем автоматического управления. В 2 т. Т 1. Вероятностные и статистические характеристики воздействий и процессов. Линейные стационарные и нестационарные системы. / Н.А. Лившиц, В.Н. Пугачёв. – Москва.: Изд. «Советское радио», 1963. – 896 с.

4. Стратонович Р.Л. Теория информации. / Р.Л. Стратонович. – Москва.: «Сов. Радио»,1975. – 424 с.

5. Дунин-Барковский И.В. Теория вероятностей и математическая статистика в технике (общая часть). / И.В. Дунин-Барковский и Н.В. Смирнов. – Москва: Госуд. изд-во технико-теоретической литературы, 1955. – 556 с.

6. Vasicek O. A test for normality based on sample entropy. / O. Vasicek // Journal of the Royal Statistical Society. Series B (Methodological). – 1976. – Vol. 38, No. 1. – pp. 54-59.

7. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников / А.И. Кобзарь – Москва: ФИЗМАТЛИТ, 2006. – 816 с.

8. Блинов П.Ю. О критериях проверки равномерности, использующих оценки энтропии. / Блинов П.Ю., Лемешко Б.Ю. // Обработка информации и моделирование. Материалы конференции. – Новосибирск.: Изд. ФГОБУ ВПО «Сибирский государственный университет телекоммуникаций и информатики», 2015. – С. 32 – 41.

9. Блинов П.Ю. О критериях проверки отклонения распределения от равномерного. / Блинов П.Ю., Лемешко Б.Ю. // Обработка информации и моделирование. Материалы конференции. – Новосибирск.: Изд. ФГОБУ ВПО «Сибирский государственный университет телекоммуникаций и информатики», 2015. – С. 31– 31.

10. Michlowicz J. V. Handbook of DIFFERENTIAL ENTROPY / J.V. Michlowicz, J.M. Nichols, Bucholtz F. – New York.: A. CHAPMAN & HALL, 2014. – 220 p.

11. Вадзинский Р.Н. Справочник по вероятностным распределениям. / Р.Н. Вадзинский. – М.: НАУКА, 2001. – 295 с.

12. Дьяконов В.П. Система компьютерной алгебры DERIVE: Самоучитель и руководство пользователя. / Дьяконов В.П. – Москва: СОЛОН-Р, 2002, – 320 с.

Надійшла до редколегії 23.02.2018

Рецензент: д-р техн. наук, проф. О.О. Можаєв, Національний технічний університет «ХПІ», Харків.

ОБЧИСЛЕННЯ ЕНТРОПІЇ ФУНКЦІЇ КВАНТИЛЕЙ (ЕНТРОПІЇ ВАСИЧЕКА)

В.Ю. Дубницький, О.Є. Петренко, О.І. Ходирєв

Поставлено задачу визначення чисельного значення ентропії функції квантилей (ентропії Васичека) для неперервної випадкової величини. Для показникового розподілу, розподілу Чампернауна, логістичного розподілу, розподілу арксинусу, розподілу Релея, розподілу Парето, розподілу Лапласа отримано в явному вигляді вирази для визначення чисельних значень ентропії функції квантилей (ентропії Васичека). Для розподілу Коши, розподілу мінімального значення, розподілу максимального значення, подвійного показникового розподілу і розподілу Вейбулла викладено процедуру розв'язання задачі чисельними методами. Запропоновано одиницю вимірювання ентропії Васичека – віт.

Ключові слова: *Васичек, ентропія, показниковий розподіл, розподіл Чампернауна, логістичний розподіл, розподіл арксинуса, розподіл Релея, розподіл Парето, розподіл Лапласа, розподіл Коши, розподіл мінімального значення, розподіл максимального значення, подвійний показниковий розподіл, розподіл Вейбулла, чисельні методи.*

THE ENTROPY CALCULATION OF FUNCTION QUANTILE (VASICEK ENTROPY)

V.Yu. Dubnitskiy, L.D. Filatova, A.I. Khodyrev

The subject is to denote the mining of entropy function quantile (Vasicek entropy). There are the expressions Exponentional Distribution, Champernown Distribution, Logistic Distribution, Pareto Distribution, Laplace Distribution, Caushy Distribution have been received in an explicit form the expressions of entropy function quantile (Vasicek entropy). The procedure of solving tasks is represented by using numerical method for Caushy Distribution, Minimum Value Distribution, Maximum Value Distribution, Double Exponentional Distribution, and Distribution Weibull. The Vasicek entropy unit is suggested.

Keywords: *entropy, quantil, Vasicek, Exponentional Distribution, Champernown Distribution, Logistic Distribution, Arcsine Distribution, Logistic Distribution, Pareto Distribution, Laplace Distribution, Caushy Distribution, Minimum Value Distribution, Maximum Value Distribution, Double Exponentional Distribution, Distribution Weibull.*