Australian Journal of Basic and Applied Sciences, 2(4): 1373-1377, 2008
ISSN 1991-8178

Asymptotic Distribution of UMVUE of *r*th Central Moments in Bernoulli Distribution

Narges Abbasi

Department of Statistics, Payame Noor University, Shiraz Centre 71365-944, Shiraz, Iran.

Abstract: In theory of U-statistics, for determine of asymptotic distribution of them, two theorems are considered. Under certain conditions, the distribution of U-statistics tend to normal distribution or linear combination of independent chi-square random variables. In this article, by a sample from Bernoulli distribution with p parameter, we determine asymptotic distribution of U-statistic for r th central moment.

key words: Asymptotic distribution, Central moments, Point estimation, U-Statistic.

INTRODUCTION

In the theory of U-statistics, we consider a functional θ defined on a set F distribution functions on R: $\theta = \theta(F)\ F \in$ F. The $\theta \in \theta(F)$ estimated by using a sample from the random variables $X_1, X_2, \ldots, X_n$ which are independently and identically distributed with distribution function F. Halmos (1946). proved that the functional θ admits an unbiased estimator if and only if there is a function h of k variables such that

$$\theta = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h(x_1, \ldots, x_k)\, Fd(x_1) Fd(x_2) \ldots Fd(x_n) \tag{1}$$

A functional satisfying equation (1) for some function h is called a *regular statistical functional* of degree k, and the function h is called the *kernel* of the functional. If a functional can be written as a regular statistical functional then optimal unbiased estimators can be constructed.

For a distribution function F let $\mu'_r = \int_{-\infty}^{+\infty} x^r dF$, the r th moment about 0, and let $\mu_r = \int_{-\infty}^{+\infty} (x - \mu'_1)^r dF$, the rth central moment or

$$\mu_r = \sum_{i=0}^{r-2} \binom{r}{i} (-1)^i \mu_1^i E(X^{r-i}) + (-1)^{r-1}(r-1)\mu_1^r. \tag{2}$$

An unbiased estimator for μ_r can be write

$$h_r^*(X_1, \ldots, X_r) = \sum_{i=1}^{r-2} \binom{r}{i} (-1)^i X_1^{r-i} \prod_{j=2}^{i+1} X_j + (-1)^{r-1} \prod_{j=1}^{r} X_j.$$

h_r^* is not symmetric about its arguments, but it appears useful in next sections.
Heffernan, P.M. (1997). obtained an estimator of the r th central moment of a distribution, which unbiased for all distributions for which the first r moments exits. There is a unique symmetric unbiased estimator of μ_r

$$U_r(x_1, \ldots, x_n) = \frac{(n-r)!}{n!} \sum h_r(x_{i_1}, \ldots, x_{i_r}),$$

where the sum extends over all $\frac{(n-r)!}{n!}$ permutations $(i_1, \ldots, i_r)$ of r distinct integers chosen from $1, 2, \ldots, n$ and

$$h_r(x_{i_1}, \ldots, x_{i_r}) = \sum_{j=0}^{r-2} (-1)^j \frac{1}{r-j} \sum x_{i_1}^{r-j} x_{i_2} \ldots x_{i_{j+1}} + (-1)^{r-1}(r-1) x_1 \cdots x_r$$

Corresponding Author: Narges Abbasi, Department of Statistics, Payame Noor University, Shiraz Centre 71365-944, Shiraz, Iran.
E-mail: abbasi@spnu.ac.ir

where the second summation is over $i_1,\ldots,i_{j+1} = 1$ to r with $i_1 \neq i_2 \neq \ldots \neq i_{j+1}$ and $i_1 < i_2 < \cdots < i_{j+1}$

$U_r(x_1,\ldots,x_n)$ has, among all estimators which are unbiased for all F in F, minimum variance for each F in F. This follows from theorem 3 of section 1.1 of Lee(1990). U_r does not have a presentation of the role of sample moments on estimation of μ_r ,obviously.

In Bernoulli family, the problem has some changes. We know $Y = \sum^n X_i$ is sufficient statistic for p.

$$\widehat{p^m} = E\left(\prod_{i=1}^{m} X_i \mid Y = y\right) = \frac{\binom{n-m}{y-m}}{\binom{n}{y}}; \qquad y \geq m, \quad n \geq m.$$

Again, rewrite the r th central moment has the following form

$$\mu_r = \sum_{i=0}^{r-2} \binom{r}{i} (-1)^i p^{i+1} + (-1)^{r-1}(r-1)p^r.$$

So the UMVUE for μ_r equals to

$$U_{p,r} = E\widehat{(X-\mu)^r}$$
$$= \frac{1}{\binom{n}{y}} \left(\left\{ \sum_{i=0}^{r-2} \binom{r}{i} (-1)^i \binom{n-i-1}{y-i-1} \right\} + (-1)^{r-1}(r-1)\binom{n-r}{y-r} \right). \tag{3}$$

This result is the answer well-known problem whatever it can be find in most inference books [Rohatgi 1976]. But here the main rule of $U_{p,r}$ is identity with U_r and it is simplified to the formula of U_r . In special cases, we can see

$$U_{p,2} = \frac{n}{n-1} M_2,$$
$$U_{p,3} = \frac{n^2}{(n-1)(n-2)} M_3,$$

where M_2 and M_3 are 2th and 3th sample central moments respectively.

In this article, section 2 has review of history of asympototic distribution of U-statistics, and says about two theorems which we apply to obtain main result. Section 3 deals to asymptotic distribution of $U_{p,r}$.

Asymptotic distribution of U-Statistic:

Consider a symmetric kernel h satisfying
$E_F(h^2(X_1,\ldots,X_k)) < \infty.$

We shall make use of the function h_c . and $\tilde{h}_k$. $h_k = h$ and for $1 \leq c \leq k-1$,

$$h_c = E_F(h(x_1,\ldots,x_c,X_{c+1},\ldots,X_k)).$$

that $\tilde{h} = h - \theta$ $\tilde{h}_c = h_c - \theta$ Define $\xi_0 = 0$ and, for $1 \leq c \leq k-1$

$$\xi_c = \mathrm{var}_F(h_c(X_1,\ldots,X_c)).$$

The following results were established by Hoeffding (1948).

Theorem 1. If $E_F h^2 < \infty$ and $\xi_1 > 0$ then $\sqrt{n}(U_n - \theta) \xrightarrow{D} N(0, k^2\xi_1).$

For the function $\tilde{h}_2(x_1, x_2)$ associated with the kernel $h = h(x_1, \ldots, x_k) \quad (k \geq 2)$ an operator A on the function space $L_2(R, F)$ was defined by

$$A g(x) = \int_{-\infty}^{+\infty} h_2(x,y) g(y) dF(y), \quad x \in \Re, \quad g \in L_2. \tag{4}$$

That is, A takes a function g into a new function Ag. In connection with any such operator A, the associated eigenvalues $\lambda_1, \lambda_2, \ldots$ to be the real number λ (not necessarily distinct) corresponding to the distinct solutions $g_1, g_2, \ldots$ of the equation $Ag - \lambda g = 0$.

The following theorem was established (Korolyuk and Borovshich (1994)).

Theorem 2: If $E_F h^2 < \infty$ and $\xi_1 = 0 < \xi_2$, then $n(U_n - \theta) \xrightarrow{D} \frac{k(k-1)}{2} Y$ where Y is a random variable of the form $Y = \sum_{j=1}^{\infty} \lambda_j (\chi_j^2 - 1)$, where $\chi_1^2, \chi_2^2, \ldots$ are independent χ^2 variables.

The essential difference between theorem 1 and theorem 2 in their conditions is about the value of ξ_1. This point will tend to finding the group of distributions that they satisfy in conditions of theorem 2.

RESULTS AND DISCUSSION

Abbasi (2008) for r=3,4 found some distribution which they satisfied on theorem 2. But it is interesting that for even central moments for Bernoulli with $p = \frac{1}{2}$ theorem 2 always occurs. Theorem 3 states this result.

Theorem 3. If $X_1, \ldots, X_n$ are iid Bernoulli random variables, $Y = \sum_{i=1}^{n} X_i$ and $U_{p,r}$ is given in (3) formula then

$$i) \qquad \sqrt{n}(U_{p,r} - \mu_r) \to N(0, r^2 \xi_1); \quad r \in N, p \neq \frac{1}{2},$$

$$ii) \qquad \sqrt{n} U_{p,r} \to N(0, 2^{-2r-1}(r-1)^2); \quad r\ odd, p = \frac{1}{2},$$

$$i) \quad n(U_{p,r} - 2^{-r}) \to 2^{-r-1} r(r-3)(\chi_1^2 - 1); \quad r\ even, p = \frac{1}{2}.$$

Proof. The parts of i and ii will be obtained from theorem 1. Using rules of counting, and change h_r^* to $\tilde{h}_r$ as a symmetric function with respect to arguments. According to section 2

$$\tilde{h}_1(x) = \frac{1}{r} \left\{ \left[\sum_{j=0}^{r-2} \binom{r}{j} (-1)^j x^{r-j} p^j \right] + \left[x \sum_{j=0}^{r-2} \binom{r}{j} (-1)^j p^j j \right] \right.$$
$$\left. + \left[\sum_{j=0}^{r-2} \binom{r}{j} (-1)^j p^{j+1} (r-1-j) \right] \right\} + (-1)^{r-1} x p^{r-1} (r-1) - p^r.$$

Under Bernoulli distribution the variance of $\tilde{h}_1$ will be obtained

$$\xi_1 = \mathrm{var}(\tilde{h}_1(X))$$
$$= \frac{1}{p(1-p)r^2} [-(1-p)^r p^2 r + (1-p)^r p - (1-p)^r p^2$$
$$+ (-1)^r p^r r + 2(-1)^{r+1} p^{r+1} r - (-1)^r p^{r+1} + (-1)^r p^{r+2} r + (-1)^r p^{r+2}]^2; \qquad p \neq \tfrac{1}{2},$$
$$= \begin{cases} \frac{(r-1)^2}{r^2} 2^{-2r-1} & r \ \text{is odd}, \quad p = \frac{1}{2} \\ 0 & r \ \text{is even}, \quad p = \frac{1}{2} \end{cases}$$

From the above results, for r even, and $p=\frac{1}{2}$ the ξ_1 equals to zero. That is the conditions of theorem 1 was failed and we must apply to theorem 2. By similar method that for finding the $\tilde{h}_1(x)$ used more calculation to obtain $\tilde{h}_2(x,y)$

$$
\begin{aligned}
\tilde{h}_2(x,y)=\frac{1}{r(r-1)}\Bigg\{&\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j x^{r-j}p^j(r-1-j)\right]\\
&+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j y^{r-j}p^j(r-1-j)\right]\\
&+\left[\sum_{j=2}^{r-2}\binom{r}{j}(-1)^j xyp^{j-1}(j-1)\right]\\
&+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j x^{r-j}yp^{j-1}j\right]+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j y^{r-j}xp^{j-1}j\right]\\
&+\left[\sum_{j=1}^{r-2}\binom{r}{j}(-1)^j xp^j(r-1-j)\right]+\left[\sum_{j=1}^{r-2}\binom{r}{j}(-1)^j yp^j(r-1-j)\right]\\
&+\left[\sum_{j=0}^{r-3}\binom{r}{j}(-1)^j p^{j+1}(r-1-j)(r-2-j)\right]\Bigg\}+(-1)^{r-1}xyp^{r-2}(r-1)-p^r.
\end{aligned}
$$

Consider $g(x)=\sum_{i=0}^{r}a_i$ and the system of equation (4).

$$
\lambda\, g(x)=\frac{1}{2}\tilde{h}_2(x,0)\,g(0)+\frac{1}{2}\tilde{h}_2(x,1)\,g(0); \qquad x\in R,
$$

with

$$
\begin{aligned}
\tilde{h}_2(x,0)=\frac{1}{r(r-1)}\Bigg\{&\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j x^{r-j}p^j(r-1-j)\right]\\
&+\left[\sum_{j=1}^{r-2}\binom{r}{j}(-1)^j x\,p^j(r-1-j)\right]+\\
&\left[\sum_{j=0}^{r-3}\binom{r}{j}(-1)^j p^{j+1}(r-1-j)(r-2-j)\right]\Bigg\}-p^r,
\end{aligned}
$$

$$
\begin{aligned}
\tilde{h}_2(x,1)=\frac{1}{r(r-1)}\Bigg\{&\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j x^{r-j}p^j(r-1-j)\right]\\
&+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j p^j(r-1-j)\right]\\
&+\left[\sum_{j=2}^{r-2}\binom{r}{j}(-1)^j x\,p^{j-1}(j-1)\right]\\
&+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j x^{r-j}p^{j-1}j\right]+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j x\,p^{j-1}j\right]\\
&+\left[\sum_{j=1}^{r-2}\binom{r}{j}(-1)^j x\,p^j(r-1-j)\right]+\left[\sum_{j=1}^{r-2}\binom{r}{j}(-1)^j p^j(r-1-j)\right]\\
&+\left[\sum_{j=0}^{r-3}\binom{r}{j}(-1)^j p^{j+1}(r-1-j)(r-2-j)\right]\Bigg\}+(-1)^{r-1}x\,p^{r-2}(r-1)-p^r,
\end{aligned}
$$

or equivalently,

$$
\begin{aligned}
2r(r-1)\lambda a_{r-j} = &\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j p^j (r-1-j)\right]a_0 \\
&+\{\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j p^j (r-1-j)\right] \\
&+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j p^{j-1} j\right]\}\sum_{i=0}^{r} a_i; \qquad j = 0,1,\ldots,r-2.
\end{aligned}
$$

$$
\begin{aligned}
2r(r-1)\lambda a_1 = &\left[\left[\sum_{j=1}^{r-2}\binom{r}{j}(-1)^j p^j (r-1-j)\right]\right]a_0 + \{\left[\sum_{j=2}^{r-2}\binom{r}{j}(-1)^j p^{j-1}(j-1)\right] \\
&+\left[\sum_{j=0}^{r-2}\binom{r}{j}(-1)^j p^{j-1} j\right] + \left[\sum_{j=1}^{r-2}\binom{r}{j}(-1)^j p^j (r-1-j)\right] \\
&+r(r-1)(-1)^{r-1} p^{r-2}(r-1)\}\sum_{i=0}^{r} a_i
\end{aligned}
$$

Replace $p = \frac{1}{2}$. The answer of above equation is

$$\lambda = \frac{2^{-r}(r-3)}{(r-1)}.$$

This completes the proof of theorem 3.

REFERENCES

Abbasi, N., 2008. On Limiting Distributions of UMVUE of Central Moments. *World Applied Sciences Journal.* Forthcoming.

Halmos, P.R., 1946. The Theory of Unbiased Estimation. *Ann. Math. Statist.*, 17: 34-43.

Heffernan, P.M., 1996. Unbiased Estimation of Central Moments By Using U-Statistics. *Royal Statistical Society*, B, 59(4): 861-863.

Hoeffding, W., 1948. A Class of Statistics With asymptotically Normal Distribution. *Ann. Math. Statist.*, 19: 293-325.

Korolyuk, V.S. and V. Borovskich, Yu, 1994. *Theory of U-Statistics*. Kluwer Academic Publishers, Dordrecht-Boston-London.

Lee, A.J., 1990. *U-statistics: Theory and Practice*. New York, Dekker,

Rohatgi, V.K., 1976. *An Introduction to Probability Theory and Mathematics Statistics*. John Wiley and Sons,