THE LIMITING DISTRIBUTION OF A FUNCTION OF TWO INDEPENDENT RANDOM VARIABLES AND ITS STATISTICAL APPLICATION

BY

M. FISZ (WARSZAWA)

1. We consider positive random variables $\xi(\lambda)$ with finite mean value $m(\lambda)$ and variance $\sigma^2(\lambda)\neq 0$ where the parameter λ can take values belonging to an unlimited set Λ of positive numbers. Let us suppose that the distribution function $F(x,\lambda)$ of $\xi(\lambda)$ can depend not only upon λ but also upon other parameters. However the values of these parameters are the same for all considered random variables.

DEFINITION 1. *We say that the random variable $\xi(\lambda)$ converges stochastically to a number c if the following relation is satisfied for an arbitrary $\varepsilon>0$:*

(1) $$\lim_{\lambda\to\infty} P(|\xi(\lambda)-c|>\varepsilon)=0, \quad \textit{where} \quad \lambda\in\Lambda.$$

DEFINITION 2. *We say that the random variable $\xi(\lambda)$ is asymptotically normal $N[u(\lambda);v(\lambda)]$ if there exist such functions $u(\lambda)$ and $v(\lambda)>0$ that for each real x the following equality holds:*

(2) $$\lim_{\lambda\to\infty} P\left(\frac{\xi(\lambda)-u(\lambda)}{v(\lambda)}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-x^2/2}\,dx=\Phi(x),$$

where $\lambda\in\Lambda$.

Let Λ_1 and Λ_2 be some subsets of Λ and let $\xi_1(\lambda_1)$ and $\xi_2(\lambda_2)$ be two independent random variables where $\lambda_1\in\Lambda_1$, $\lambda_2\in\Lambda_2$. We shall denote, for $r=1,2$,

$$\xi_r=\xi_r(\lambda_r), \qquad m_r=E(\xi_r), \qquad \sigma_r^2=D^2(\xi_r), \qquad \psi=\sqrt{\sigma_1^2+\sigma_2^2}, \qquad \beta_r=\frac{\sigma_r}{\psi}.$$

The following theorem[1]) will be proved:

THEOREM. *If*

(a) *the variable $\xi(\lambda)/m(\lambda)$ is stochastically convergent to the number 1,*

(b) *the variable $\xi(\lambda)$ is asymptotically normal $N[m(\lambda);\sigma(\lambda)]$,*

[1]) The assumptions of this theorem have been formulated by the author in a discussion with A. Rényi and J. Łukaszewicz.

(c) *the variables ξ_1 and ξ_2 are independent and the following relation holds:*

(3) $$\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}} \frac{m_2}{m_1}=1,$$

then the variable

(4) $$\zeta(\lambda_1,\lambda_2)=\frac{\xi_2-\xi_1}{(\xi_2+\xi_1)^p},$$

where p is an arbitrary positive number, is asymptotically normal

$$N\left[\frac{m_2-m_1}{(m_1+m_2)^p};\frac{\psi}{(m_1+m_2)^p}\right], \quad \textit{when} \quad \lambda_1\to\infty,\ \lambda_2\to\infty.$$

In the formula (4) we understand by $(\xi_1+\xi_2)^p$ the real and positive value of this expression.

Proof. In the first place we shall prove two lemmata.

LEMMA 1. *If $\xi(\lambda)/m(\lambda)$ is stochastically convergent to the number 1, then*

(5) $$\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}} P\left(\left|\frac{\xi_1+\xi_2}{m_1+m_2}-1\right|>\varepsilon\right)=0,$$

where $\varepsilon>0$ is an arbitrary positive number.

Proof of lemma 1. From the assumption of the lemma it follows that for sufficiently large values of λ_1 and λ_2 the probability of the occurrence of each of the following two inequalities:

(6) $$m_r(1-\varepsilon)<\xi_r<m_r(1+\varepsilon) \qquad (r=1,2)$$

is greater than $1-\delta$, where $\delta>0$ is a given arbitrarily small number. The probability of the occurrence of both inequalities (6) is — in view of the independence of ξ_1 and ξ_2 — greater than $(1-\delta)^2$. As the probability of the occurrence of the relation

(7) $$(m_1+m_2)(1-\varepsilon)<\xi_1+\xi_2<(m_1+m_2)(1+\varepsilon)$$

is at least equal to the probability of the occurrence of both inequalities (6), and δ can be arbitrarily small, the assertion of the lemma is proved.

LEMMA 2. *If $\xi(\lambda)$ is asymptotically normal $N[m(\lambda);\sigma(\lambda)]$ then the variable $\xi_2-\xi_1$ is asymptotically normal $N[m_2-m_1;\psi]$ when $\lambda_1\to\infty,\lambda_2\to\infty$.*

Proof of lemma 2. The assumption of the lemma is equivalent to the relation

$$\lim_{\lambda\to\infty}\left[e^{-m(\lambda)it/\sigma(\lambda)}\varphi\left(\frac{t}{\sigma(\lambda)},\lambda\right)-e^{-t^2/2}\right]=0, \tag{8}$$

where $\varphi(t,\lambda)$ is the characteristic function of the variable $\xi(\lambda)$ and the convergence is uniform in each finite interval of t. Let us denote

$$\eta(\lambda_1,\lambda_2)=\frac{\xi_2-\xi_1-(m_2-m_1)}{\psi}. \tag{9}$$

The characteristic function $\varphi_1(t,\lambda_1,\lambda_2)$ of $\eta(\lambda_1,\lambda_2)$ is of the form

$$\varphi_1(t,\lambda_1,\lambda_2)=e^{-m_2it/\psi}\varphi\left(\frac{t}{\psi},\lambda_2\right)e^{m_1it/\psi}\varphi\left(\frac{-t}{\psi},\lambda_1\right) \tag{10}$$

$$=A_1(t,\lambda_1,\lambda_2)A_2(t,\lambda_1,\lambda_2)=A_1A_2,$$

where

$$A_1=A_1(t,\lambda_1,\lambda_2)=e^{m_1it/\psi}\varphi\left(-\frac{t}{\psi},\lambda_1\right)=e^{m_1\beta_1it/\sigma_1}\varphi\left(-\frac{t}{\sigma_1}\beta_1,\lambda_1\right)$$

$$=\overline{e^{-m_1\beta_1it/\sigma_1}\varphi\left(\frac{t}{\sigma_1}\beta_1,\lambda_1\right)},$$

$$A_2=A_2(t,\lambda_1,\lambda_2)=e^{-m_2it/\psi}\varphi\left(\frac{t}{\psi},\lambda_2\right)=e^{-m_2\beta_2it/\sigma_2}\varphi\left(\frac{t}{\sigma_2}\beta_2,\lambda_2\right).$$

Let us denote

$$B_r=e^{-\beta_r^2t^2/2}\qquad (r=1,2).$$

As the convergence in the relation (8) is uniform in each finite interval and $0<\beta_r<1$, it follows from (8) that the following relation holds:

$$\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}}(A_r-B_r)=0\qquad (r=1,2) \tag{11}$$

and thus

$$\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}}[(A_1-B_1)(A_2-B_2)]=$$

$$=\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}}[A_1A_2-B_1B_2-B_2(A_1-B_1)-B_1(A_2-B_2)]=0.$$

Taking into account the relations (11) and (10) we obtain

$$\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}}(A_1A_2-B_1B_2)=\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}}[\varphi_1(t_1,\lambda_1,\lambda_2)-e^{-(\beta_1^2+\beta_2^2)t^2/2}] \tag{12}$$

$$=\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}}[\varphi_1(t,\lambda_1,\lambda_2)-e^{-t^2/2}]=0.$$

The assertion of Lemma 2 follows immediately from the last relation.

Let us further denote

$$\tau(\lambda_1,\lambda_2)=\frac{\dfrac{\zeta(\lambda_1,\lambda_2)-\dfrac{m_2-m_1}{(m_2+m_1)^p}}{\psi}}{(m_1+m_2)^p} \tag{13}$$

$$=\frac{\dfrac{\xi_2-\xi_1-(m_2-m_1)}{\psi}+\dfrac{m_2-m_1}{\psi}\left[1-\left(\dfrac{\xi_1+\xi_2}{m_1+m_2}\right)^p\right]}{\left(\dfrac{\xi_1+\xi_2}{m_1+m_2}\right)^p}$$

$$=\frac{\eta(\lambda_1,\lambda_2)+\dfrac{m_2-m_1}{\psi}\left[1-\left(\dfrac{\xi_1+\xi_2}{m_1+m_2}\right)^p\right]}{\left(\dfrac{\xi_1+\xi_2}{m_1+m_2}\right)^p},$$

where $\eta(\lambda_1,\lambda_2)$ is given by the formula (9).

The assumption (a) implies — in view of Lemma 1 — the relation (5), which is equivalent to the relation

$$\lim_{\substack{\lambda_1\to\infty\\ \lambda_2\to\infty}}P\left[\left|\left(\frac{\xi_1+\xi_2}{m_1+m_2}\right)^p-1\right|>\varepsilon\right]=0. \tag{14}$$

Let us now consider the expression on the right side of formula (13). From (14) it follows that its denominator converges stochastically to the number 1. The assumption (b) and Lemma 2 imply that the distribution function of the variable $\eta(\lambda_1,\lambda_2)$ converges to the function $\Phi(x)$, defined by (2). Let us further write

$$z=\frac{m_2-m_1}{\psi}(1-y^p)=\frac{m_2-m_1}{\psi}(p+\vartheta)(1-y)=-\frac{m_2-m_1}{m_1+m_2}(p+\vartheta)\varrho(\lambda_1,\lambda_2),$$

where

$$y=\frac{\xi_1+\xi_2}{m_1+m_2},\qquad \varrho(\lambda_1,\lambda_2)=\frac{\xi_1+\xi_2-(m_1+m_2)}{\psi},$$

is a function of y and $\vartheta\to 0$ when $y\to 1$. It is easily seen that ϑ converges stochastically to 0. Let $\varepsilon>0$ and $\delta>0$ be arbitrary given numbers. We have

$$P(|z|>\varepsilon)=P(|z|>\varepsilon\,\big|\,|\vartheta|>\delta)P(|\vartheta|>\delta)+P(|z|>\varepsilon\,\big|\,|\vartheta|<\delta)P(|\vartheta|<\delta);$$

$$P(|z|>\varepsilon\,\big|\,|\vartheta|<\delta)\leqslant P\left(\left|\frac{m_2-m_1}{m_1+m_2}\right|(p+|\vartheta|)|\varrho(\lambda_1,\lambda_2)|>\varepsilon\,\big|\,|\vartheta|<\delta\right)$$

$$\leqslant P\left(\left|\frac{m_2-m_1}{m_1+m_2}\right|(p+\delta)|\varrho(\lambda_1,\lambda_2)|>\varepsilon\,\big|\,|\vartheta|<\delta\right).$$

From the stochastical convergence of ϑ to 0 follows the relation

$$P(|z|>\varepsilon\,\big|\,|\vartheta|>\delta)P(|\vartheta|>\delta)\to 0. \tag{15}$$

Since $\varrho(\lambda_1,\lambda_2)$ is — as can be shown by a slight modification of the proof of Lemma 2 — asymptotically normal $N(0;1)$, the assumption (c) implies the relation

$$P\left(\left|\frac{m_2-m_1}{m_1+m_2}\right|(p+\delta)|\varrho(\lambda_1,\lambda_2)|>\varepsilon\,\big|\,|\vartheta|<\delta\right)\to 0. \tag{15'}$$

From the last two relations it follows that z is stochastically convergent to the number 0.

A well known theorem given by Cramér[2]) implies that the distribution function of the random variable $\tau(\lambda_1,\lambda_2)$ converges to $\Phi(x)$. Our theorem is thus proved.

2. We shall give some examples of the application of the theorem proved.

Example 1. The random variable $\xi(\lambda)$ has a Gamma distribution whose characteristic function is

$$\varphi(t,\lambda)=\left(1-\frac{it}{b}\right)^{-\lambda}, \tag{16}$$

where $\lambda>0$, $b>0$. One can easily find

$$m(\lambda)=\frac{\lambda}{b},\qquad \sigma^2(\lambda)=\frac{\lambda}{b^2}.$$

[2]) H. Cramér, *Mathematical methods of statistics*, Princeton 1946, § 20.6.

We have

$$\lim_{\lambda\to\infty}\varphi\left[\frac{t}{m(\lambda)},\lambda\right]=\lim_{\lambda\to\infty}\left(1-\frac{it}{\lambda}\right)^{-\lambda}=e^{it}, \tag{17}$$

$$\log\left[e^{-m(\lambda)it/\sigma(\lambda)}\varphi\left(\frac{t}{\sigma(\lambda)},\lambda\right)\right]=-\sqrt{\lambda}\,it-\lambda\log\left(1-\frac{it}{\sqrt{\lambda}}\right)=-\frac{t^2}{2}+\lambda o\left(\frac{t^2}{\lambda}\right).$$

From the last formula follows

$$\lim_{\lambda\to\infty}\left[e^{-m(\lambda)it/\sigma(\lambda)}\varphi\left(\frac{t}{\sigma(\lambda)},\lambda\right)\right]=e^{-t^2/2}. \tag{18}$$

The assumptions (a) and (b) of our theorem are thus satisfied. Consequently if the assumption (c) is satisfied, the variable $\zeta(\lambda_1,\lambda_2)$ defined by (4), is for $\lambda_1\to\infty$, $\lambda_2\to\infty$ asymptotically normal

$$N\left[\frac{\lambda_2-\lambda_1}{(\lambda_2+\lambda_1)^p}b^{p-1};\frac{1}{(\lambda_1+\lambda_2)^{p-0,5}}b^{p-1}\right].$$

Let us consider the variable

$$ns^2=\sum_{k=1}^{n}x_k^2-\frac{1}{n}\left(\sum_{k=1}^{n}x_k\right)^2,$$

where x_k $(k=1,2,\ldots,n)$ are independent and equally normally distributed $N(m;\sigma)$. The variable ns^2 is a special case of a variable distributed according to (16) for $\lambda=(n-1)/2$, $b=1/2\sigma^2$. Thus if the assumption (c) is satisfied, the variable $\zeta(n_1,n_2)$, defined by the formula

$$\zeta(n_1,n_2)=\frac{n_2s_2^2-n_1s_1^2}{(n_1s_1^2+n_2s_2^2)^p} \tag{19}$$

where s_1 and s_2 are independent, is asymptotically normal

$$N\left[\frac{n_2-n_1}{(n_1+n_2-2)^p}\,\frac{1}{(\sigma^2)^{p-1}};\frac{\sqrt{2}}{(n_1+n_2-2)^{p-0,5}}\,\frac{1}{(\sigma^2)^{p-1}}\right].$$

Let us suppose that the variances σ_1^2 and σ_2^2 in two normal populations are not known and we want to test the hypothesis $H_0(\sigma_1=\sigma_2)$. We then find $n_1s_1^2$ and $n_2s_2^2$ from two samples chosen from the considered populations, where n_1 and n_2 are the numbers of elements in the samples. Thus if the hypothesis $H_0(\sigma_1=\sigma_2)$ is true and n_1 and n_2 are sufficiently large and the quotient n_2/n_1 is near 1, the variable $\zeta(n_1,n_2)$ defined by (19), with $p=1$ is asymptotically normal

$$N\left[\frac{n_2-n_1}{n_1+n_2-2};\sqrt{\frac{2}{n_1+n_2-2}}\right].$$

This distribution is independent of σ and the hypothesis $H_0(\sigma_1=\sigma_2)$ can be tested. For $n_1=n_2=n$ we have

$$\zeta(n,n)=\frac{s_2^2-s_1^2}{s_1^2+s_2^2}. \tag{20}$$

The variable $\zeta(n,n)\sqrt{n-1}$ is asymptotically normal $N(0;1)$[3].

Table 1 shows that the convergence of the distribution function of $\zeta(n,n)\sqrt{n-1}$ to the function $\Phi(x)$ given by (2) is very rapid. This table has been constructed in the following way: For some values of γ the values of x_γ have been found from the relation

$$P(x>x_\gamma)=\gamma$$

where x is normally distributed $N(0;1)$. For some values of n, the values of α, satisfying the equality

$$\alpha=P[\zeta(n,n)\sqrt{n-1}>x_\gamma]=\int_t^1\frac{u^{(n-3)/2}(1-u)^{(n-3)/2}}{B\left(\frac{n-1}{2},\frac{n-1}{2}\right)}du$$

where

$$t=\frac{1+x_\gamma/\sqrt{n-1}}{2},$$

have been found from Pearson's tables[4].

TABLE 1

γ	α			
	$n=5$	$n=10$	$n=17$	$n=37$
0,400	0,406	0,403	0,400	0,401
0,300	0,308	0,304	0,302	0,300
0,200	0,204	0,202	0,202	0,201
0,150	0,146	0,147	0,150	0,151
0,100	0,086	0,098	0,098	0,099
0,050	0,022	0,041	0,045	0,048
0,025	—	0,015	0,020	0,023
0,010	—	0,002	0,006	0,008
0,005	—	0,001	0,002	0,004

From table 1 we see that the approximation is remarkably good even for small n.

[3]) This fact can also be deduced from the asymptotical normality of the Beta distribution.

[4]) K. Pearson, *Tables of the incomplete Beta-function*, Cambridge 1934.

Let us observe that it is convenient to take in (19) $n_1=n_2$ and $p=0{,}5$ if we want to test hypothesis $H_0(\sigma_1=\sigma_2=\sigma)$, where σ is a specified number. The variable $\zeta(n,n)$ is then asymptotically normal $N(0;\sigma\sqrt{2})$.

Example 2. $\xi(\lambda)$ is a Poisson variable whose characteristic function is

$$\varphi(t,\lambda)=e^{\lambda(e^{it}-1)}.$$

We have here

$$m(\lambda)=\sigma^2(\lambda)=\lambda.$$

It can easily be found that the assumptions (a) and (b) are satisfied. Thus if the assumption (c) is satisfied, the variable $\zeta(\lambda_1,\lambda_2)$ is asymptotically normal

$$N\left[\frac{\lambda_2-\lambda_1}{(\lambda_1+\lambda_2)^p},\frac{1}{(\lambda_1+\lambda_2)^{p-0,5}}\right].$$

Consequently if $\lambda_1=\lambda_2=\lambda$, the limiting distribution is $N[0;1/(2\lambda)^{p-1/2}]$.

Let us suppose that the values of the parameter λ of two independent Poisson variables are not known. Then the hypothesis $H_0(\lambda_1=\lambda_2)$ can be tested by observing the value of the variable

$$\zeta(\lambda_1,\lambda_2)=\frac{\xi_2-\xi_1}{\sqrt{\xi_1+\xi_2}}$$

which — if the hypothesis $H_0(\lambda_1=\lambda_2)$ is true — is asymptotically normal[5]) $N(0;1)$. On the other hand the hypothesis $H_1(\lambda_1=\lambda_2=\lambda)$, where λ is some specified number, can be tested by observing the variable

$$\zeta(\lambda_1,\lambda_2)=\frac{\xi_2-\xi_1}{\xi_1+\xi_2}$$

which — if the hypothesis H_1 is true — is asymptotically normal $N(0;1/\sqrt{2\lambda})$.

One can give many other examples of the application of the theorem proved. We shall limit ourselves to mention further two examples without going into details.

Example 3. Two independent random variables $\xi_r(n_r)$ $(r=1,2)$ are binomially distributed:

$$P[\xi_r(n_r)=j]=\binom{n_r}{j}w_r^j(1-w_r)^{n_r-j},$$

[5]) This fact, surmised by J. Oderfeld, was the starting point of the problem dealt with by the author.

where w_1 and w_2 are unknown. The hypothesis $H_0(w_1=w_2=w)$, where w is a specified number, can be tested by observing the variable $\zeta(n_1,n_2)$, defined by (4), which — if the assumption (c) and the hypothesis H_0 are satisfied — is asymptotically normal

$$N\left[\frac{n_2-n_1}{(n_1+n_2)^p}\,w^{1-p};\ \frac{\sqrt{1-w}}{(n_1+n_2)^{p-0,5}}\,w^{0,5-p}\right].$$

EXAMPLE 4. The variable $\xi(\lambda)$ is distributed according to the negative binomial law, given by the formula

$$P[\xi(\lambda)=j]=(-1)^j\binom{-\lambda}{j}w^j(1-w)^\lambda,$$

where $j=0,1,2,\ldots,\ \lambda>0,\ 0<w<1$. Here we have

$$m(\lambda)=\frac{w\lambda}{1-w};\qquad \sigma^2(\lambda)=\frac{w\lambda}{(1-w)^2}.$$

The variable $\zeta(\lambda_1,\lambda_2)$ defined by (4) — if the assumption (c) is satisfied — is asymptotically normal

$$N\left[\frac{\lambda_2-\lambda_1}{(\lambda_1+\lambda_2)^p}\,\frac{w^{1-p}}{(1-w)^{1-p}};\ \frac{[(\lambda_1+\lambda_2)w]^{0,5-p}}{(1-w)^{1-p}}\right].$$

Our theorem can thus be applied in particular to testing parametric hypothesis concerning Pascal variables since they have a negative binomial distribution with an integer value λ.

STATISTICAL ESTIMATION OF PARAMETERS IN MARKOV PROCESSES

BY

O. LANGE (WARSZAWA)

1. Methods of estimation. Consider a simple Markov process with the transition function

(1.1) $$f(t_0,x_0;t_k,x_k;\Theta_1,\Theta_2,\ldots).$$

The transition function expresses the conditional probability (for discrete processes), or the conditional probability density (for continuous processes), that the random variable $\xi(t)$ will assume the value x_k at the moment t_k if its value is x_0 at the moment t_0. This function contains certain parameters $\Theta_1,\Theta_2,\ldots$ the values of which have to be determined from statistical observation.

In Markov processes this can be done by the method of maximum likelihood, which consists in choosing the estimators of the parameters $\Theta_1,\Theta_2,\ldots$ so as to maximize the probability or probability density of an observed set of realizations of the stochastic process. The method of maximum likelihood can be applied in several ways.

If the realizations of the stochastic process can be repeated many times (as, for instance, in the laboratory or in industrial production) we take n independent realizations of the process and perform on each realization a pair of observations at the moments, say, $t_0^{(r)}$ and $t_k^{(r)}$. The superscript r stands for the r-th realization $(r=1,2,\ldots,n)$. Denote by $x_i^{(r)}$ the result of the observation carried out on the r-th realization at the moment $t_i^{(r)}$, where $i=0,k$. Since the pairs of observations are independent, their likelihood function is

(1.2) $$L_1=\prod_{r=1}^{n}f(t_0^{(r)},x_0^{(r)};t_k^{(r)},x_k^{(r)};\Theta_1,\Theta_2,\ldots).$$

The estimators of $\Theta_1,\Theta_2,\ldots$, which will be denoted by $\hat\Theta_1,\hat\Theta_2,\ldots$, are determined from the condition $L_1=\max$.

This way of using the method of maximum likelihood will be called *cross section estimation*, or *space estimation* (over the space of realizations of the process). The estimators thus derived will be called *cross section* or *space estimators*.