REFERENCES

(1) LEFSCHETZ, S. Correspondences between algebraic curves. *Ann. Math.* (2), 28 (1927), 342–54.
(2) SCOTT, D. B. Invariant groups associated with an algebraic surface. *Proc. Cambridge Phil. Soc.* 36 (1940), 414–23.
(3) SCOTT, D. B. Point-curve correspondences. I. *Proc. Cambridge Phil. Soc.* 41 (1945), 135–45.
(4) SEVERI, F. Le corrispondenze fra i punti di una curva variabile in un sistema lineare sopra una superficie algebriche. *Math. Ann.* 74 (1913), 515–44.

QUEEN MARY COLLEGE
UNIVERSITY OF LONDON

THE LARGE-SAMPLE THEORY OF SEQUENTIAL TESTS

BY M. S. BARTLETT

Received 25 October 1945

1. The general theory underlying sampling inspection methods introduced during the war under the name of sequential analysis, applicable in cases where the sampling units can be taken serially, has mainly been developed by Wald and is now published (5). The original purpose of the present investigation was to exhibit the main structure of the distributional theory relating to the size of sample required by noting its relationship with the classical 'random-walk' problem. Thus in Part I of this paper are derived the distribution and characteristic functions of the absorption time for one-dimensional random-walk theory with constant drift and absorbing barriers. Attention is confined to the asymptotic case of numerous independent displacements, for which it is well known that the unrestricted total displacement becomes Gaussian; the motion may also be treated as continuous. Since these results were obtained Wald has independently given an alternative and more general discussion of this problem ((4); see also Tweedie (3)), but the extension here of the method of images used by Chandrasekhar (1) still appears of interest, and has advantages over alternative direct methods of solution for the distribution function involving Fourier expansion (cf. 2).

In Part II these results are applied to sequential analysis, and in particular is shown the relation of this large-sample theory to Fisher's information function.

PART I

One-dimensional random-walk theory with constant drift and absorbing barriers

2. For one-dimensional diffusion from a point source with a constant drift in the random displacement X at rate m (>0) per unit time, and a growth in variance at rate σ^2, the differential equation for the cumulant function $K\{iu\} \equiv \log E\{e^{ixu}\}$ (where E denotes average value) will be under the conditions noted in § 1 above

$$\frac{\partial K}{\partial t} = miu - \tfrac{1}{2}\sigma^2 u^2, \tag{1}$$

https://doi.org/10.1017/S0305004100022994 Published online by Cambridge University Press

the corresponding equation for the probability density function $f(x)$ being readily obtained by Fourier inversion as

$$\frac{\partial f}{\partial t}+m\frac{\partial f}{\partial x}=\tfrac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}. \tag{2}$$

Since (1) may be written $$\frac{\partial K}{\partial t}=\tfrac{1}{2}\sigma^2\left(iu+\frac{m}{\sigma^2}\right)^2-\frac{1}{2}\frac{m^2}{\sigma^2},$$

the solution of (2) may be obtained from the simpler form when $m=0$ by writing the function $K\{iu\}$ in the form

$$K\{iu\}=K'\left\{iu+\frac{m}{\sigma^2}\right\}-\frac{1}{2}\frac{m^2t}{\sigma^2},$$

or correspondingly $$f=(ge^{mx/\sigma^2})\,e^{-\frac{1}{2}m^2t/\sigma^2},$$

where g satisfies the equation $$\frac{\partial g}{\partial t}=\tfrac{1}{2}\sigma^2\frac{\partial^2 g}{\partial x^2}. \tag{3}$$

If an absorbing barrier exists at $x=a$ (>0), the solution of (3) is known to be (1, 2)

$$\left.\begin{aligned} g&=(2\pi\sigma^2t)^{-\frac{1}{2}}\{e^{-\frac{1}{2}x^2/\sigma^2t}-e^{-\frac{1}{2}(2a-x)^2/\sigma^2t}\} && (x\leqslant a),\\ &=0 && (x>a),\end{aligned}\right\} \tag{4}$$

the solution being equivalent to subtracting from the standard diffusion term g_0 that due to an 'image' source at $x=2a$, say g_{2a}.

If two absorbing barriers exist at $x=+a$ and $x=-b$ $(b>0)$, we shall similarly obtain a solution in terms of the infinite series of images at $2a$, $2a+2b$, $4a+2b$, ..., $-2b$, $-2a-2b$, $-2a-4b$, ...,

$$\left.\begin{aligned} g&=g_0-g_{2a}-g_{-2b}+g_{2a+2b}+g_{-2a-2b}-g_{4a+2b}-g_{-2a-4b}\cdots && (-b\leqslant x\leqslant a),\\ &=0 && (x<-b,\ x>a).\end{aligned}\right\} \tag{5}$$

Returning to the general solution f of the full equation (2), we note that the distribution in time $h(t)$ at which absorption takes place is determined by the relation

$$h(t)=-\frac{\partial}{\partial t}\int_{-b}^{a}f(x,t)\,dx,$$

since $$\int_0^t h(t)\,dt=1-\int_{-b}^{a}f(x,t)\,dx, \tag{6}$$

if $f(x,0)$ is normalized. We have further

$$\int_0^\infty h(t)\,dt=\left[-\int_{-b}^{a}f(x,t)\,dx\right]_{t=0}^{t=\infty}=1, \tag{7}$$

provided that the upper limit $t=\infty$ vanishes.

Characteristic function of $h(t)$

3. (i) *Case $b=-\infty$.* We have in the simple case of one barrier (i.e. $b=-\infty$)

$$h(t)=-\frac{\partial}{\partial t}\left\{F\left(\frac{a-mt}{\sigma\sqrt{t}}\right)-\exp\left(\frac{2am}{\sigma^2}\right)F\left(\frac{-a-mt}{\sigma\sqrt{t}}\right)\right\}, \tag{8}$$

where $$F(y)=\int_{-\infty}^{y}\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}v^2}dy.$$

https://doi.org/10.1017/S0305004100022994 Published online by Cambridge University Press

Hence we obtain after differentiation,

$$h(t) = \frac{a}{\sqrt{(2\pi)\sigma t^{\frac{3}{2}}}} \exp\left\{-\frac{1}{2}\frac{(a-mt)^2}{\sigma^2 t}\right\}. \tag{9}$$

If $\frac{mt}{a} = \tau$, $\sqrt{\left(\frac{am}{\sigma^2}\right)} = \alpha$ this becomes

$$h(t)\,dt = \frac{\alpha\,d\tau}{\sqrt{(2\pi)\,\tau^{\frac{3}{2}}}} \exp\{-\tfrac{1}{2}\alpha^2(\tau+\tau^{-1}-2)\}, \tag{10}$$

and if, further, $T = \tau/\alpha^2$,

$$h(t)\,dt = \frac{dT}{\sqrt{(2\pi)\,T^{\frac{3}{2}}}} \exp\{-\tfrac{1}{2}(\alpha^4 T + T^{-1} - 2\alpha^2)\}. \tag{11}$$

From (11) we note with the use of relation (7) that

$$\int_0^\infty \frac{dT}{\sqrt{(2\pi)\,T^{\frac{3}{2}}}} e^{-\frac{1}{2}(\alpha^4 T + T^{-1})} = e^{-\alpha^2}. \tag{12}$$

Hence for the characteristic function $M(\theta) = E(e^{i\theta t})$ we obtain

$$\begin{aligned} M(\theta) = E(e^{ai\theta\tau/m}) &= e^{\alpha^2} \int_0^\infty \frac{dT}{\sqrt{(2\pi)\,T^{\frac{3}{2}}}} e^{-\frac{1}{2}\{T^{-1} + T(\alpha^4 - 2\alpha^2 ai\theta/m)\}}, \\ &= \exp\{\alpha^2 - (\alpha^4 - 2\alpha^2 ai\theta/m)^{\frac{1}{2}}\}, \end{aligned} \tag{13}$$

and

$$K(\theta) = \alpha^2 - (\alpha^4 - 2\alpha^2 ai\theta/m)^{\frac{1}{2}}, \tag{14}$$

whence

$$\left.\begin{aligned} \kappa_1 &= E(t) = \frac{a}{m}, \\ \kappa_2 &= \sigma^2(t) = \left(\frac{a}{m}\right)^2 \Big/ \alpha^2 = \frac{a\sigma^2}{m^3}, \\ \kappa_3 &= 3\left(\frac{a}{m}\right)^3 \Big/ \alpha^4 = 3\frac{a\sigma^4}{m^5}, \text{ etc.} \end{aligned}\right\} \tag{15}$$

(ii) *Case b finite.* While the above solution provides an approximate solution for the double barrier problem if $bm/\sigma^2 \gg 1$, the exact solution may also be obtained by similar methods.

We obtain from (5), and the relation between f and g,

$$\begin{aligned} f = g_{mt} &- e^{2am/\sigma^2} g_{2a+mt} - e^{-2bm/\sigma^2} g_{-2b+mt} + e^{2(a+b)m/\sigma^2} g_{2a+2b+mt} \\ &+ e^{-2(a+b)m/\sigma^2} g_{-2a-2b+mt} - e^{2(2a+b)m/\sigma^2} g_{4a+2b+mt} - e^{-2(a+2b)m/\sigma^2} g_{-2a-4b+mt} \cdots, \end{aligned} \tag{16}$$

whence

$$\begin{aligned} \int_t^\infty h(t)\,dt = {} & F\left(\frac{a-mt}{\sigma\sqrt{t}}\right) - F\left(-\frac{b-mt}{\sigma\sqrt{t}}\right) \\ & - \exp\left(\frac{2am}{\sigma^2}\right)\left\{F\left(\frac{-a-mt}{\sigma\sqrt{t}}\right) - F\left(\frac{-2a-b-mt}{\sigma\sqrt{t}}\right)\right\} \cdots \end{aligned}$$

and

$$\begin{aligned} h(t) = h_a &- e^{-2bm/\sigma^2} h_{a+2b} + e^{-2(a+b)m/\sigma^2} h_{3a+2b} - e^{-2(a+2b)m/\sigma^2} h_{3a+4b} \cdots \\ &+ e^{-2bm/\sigma^2} h_b - e^{-2(a+b)m/\sigma^2} h_{2a+b} + e^{-2(a+2b)m/\sigma^2} h_{2a+3b} - e^{-2(2a+2b)m/\sigma^2} h_{4a+3b} \cdots \end{aligned} \tag{17}$$

where h_a is given by (9). From the correspondence with the sequence of image points, the terms above for $h(t)$ have been listed in two rows the first of which represents net absorption at the barrier at a, the second, net absorption at $-b$; i.e. the first term

https://doi.org/10.1017/S0305004100022994 Published online by Cambridge University Press

represents 'gross' absorption at a uncorrected for that at $-b$, the second term in the first row is a first correction to absorption at a due to previous (uncorrected) absorption at $-b$, etc.

Hence, writing $\Theta = 1-(1-2i\theta\sigma^2/m^2)^{\frac{1}{2}}, \quad \beta = \sqrt{(bm/\sigma^2)},$

we may write

$$M(\theta) = e^{\alpha^2\Theta} - e^{-2\beta^2} e^{(\alpha^2+2\beta^2)\Theta} + e^{-2(\alpha^2+\beta^2)} e^{(3\alpha^2+2\beta^2)\Theta} \ldots$$
$$+ e^{-2\beta^2} e^{\beta^2\Theta} - e^{-2(\alpha^2+\beta^2)} e^{(2\alpha^2+\beta^2)\Theta} + e^{-2(\alpha^2+2\beta^2)} e^{(2\alpha^2+3\beta^2)\Theta} \ldots$$

corresponding to the two rows in (17), whence

$$M(\theta) = \frac{e^{\alpha^2\Theta}(1-e^{2\beta^2(\Theta-1)})}{1-e^{2(\alpha^2+\beta^2)(\Theta-1)}} + \frac{e^{\beta^2(\Theta-2)}(1-e^{2\alpha^2(\Theta-1)})}{1-e^{2(\alpha^2+\beta^2)(\Theta-1)}}, \tag{18}$$

where the two terms correspond to net absorption at the first and second barriers.

Putting $\Theta = 1$, we have the chances of absorption at each barrier as

$$p_a = \frac{1-e^{-2\beta^2}}{1-e^{-2(\alpha^2+\beta^2)}}, \quad p_b = \frac{(1-e^{-2\alpha^2})\,e^{-2\beta^2}}{1-e^{-2(\alpha^2+\beta^2)}}. \tag{19}$$

Expanding (18) in powers of θ, we obtain further for the mean time κ_1 for absorption at either barrier

$$\frac{m^2\kappa_1}{\sigma^2} = \frac{\alpha^2(1-e^{-2\beta^2}) - \beta^2 e^{-2\beta^2}(1-e^{-2\alpha^2})}{1-e^{-2(\alpha^2+\beta^2)}} \tag{20}$$
$$= p_a\alpha^2 - p_b\beta^2.$$

To obtain the variance κ_2, we shall first write $M(\theta)$ in (18) as

$$M(\theta) = p_a M_a + p_b M_b. \tag{21}$$

Then

$$\left.\begin{aligned} \kappa_1 &= p_a(\kappa_1)_a + p_b(\kappa_1)_b, \\ \kappa_2 &= p_a\{(\kappa_2)_a + (\kappa_1)_a^2\} + p_b\{(\kappa_2)_b + (\kappa_1)_b^2\} - \kappa_1^2, \end{aligned}\right\} \tag{22}$$

where

$$K_a \equiv \log M_a = \alpha^2\Theta + \log\{1-e^{-2\beta^2} e^{2\beta^2\Theta}/(1-e^{-2\beta^2})\}$$
$$- \log\{1-e^{-2(\alpha^2+\beta^2)} e^{2(\alpha^2+\beta^2)\Theta}/(1-e^{-2(\alpha^2+\beta^2)})\}. \tag{23}$$

Hence

$$\left.\begin{aligned} \frac{m^2(\kappa_1)_a}{\sigma^2} &= \left\{\alpha^2 - \frac{2\beta^2 e^{-2\beta^2}}{1-e^{-2\beta^2}} + \frac{2(\alpha^2+\beta^2)\,e^{-2(\alpha^2+\beta^2)}}{1-e^{-2(\alpha^2+\beta^2)}}\right\}, \\ \frac{m^4(\kappa_2)_a}{\sigma^4} &= \left\{\frac{m^2(\kappa_1)_a}{\sigma^2} - \frac{4\beta^4 e^{-4\beta^2}}{(1-e^{-2\beta^2})^2} + \frac{4(\alpha^2+\beta^2)^2\,e^{-4(\alpha^2+\beta^2)}}{(1-e^{-2(\alpha^2+\beta^2)})^2}\right\}, \end{aligned}\right\} \tag{24}$$

with similar expressions for κ_b, $(\kappa_1)_b$, $(\kappa_2)_b$, etc. with α and β interchanged. We shall note further for reference that while $p_b = 1-p_a$ is not obtained from p_0 by interchanging α and β, we may define a new probability p_b' so obtained, and this will represent the probability of absorption at the barrier at b when the drift m is reversed in direction. If we solve for α^2 and β^2 in terms of p_a and p_b', we obtain

$$e^{-2\alpha^2} = p_a'/p_a, \quad e^{-2\beta^2} = p_b/p_b', \tag{25}$$

and

$$\frac{m^2\kappa_1}{\sigma^2} = \frac{1}{2}\left(p_a \log\frac{p_a}{p_a'} + p_b \log\frac{p_b}{p_b'}\right). \tag{26}$$

When the expressions p_a'/p_a, p_b/p_b' in (25) are small, we obtain as the second approximation to the result $\kappa_2 = \sigma^2\kappa_1/m^2$ in (15), by writing $(\kappa_2)_a \sim \sigma^2(\kappa_1)_a/m^2$ in (24),

$$\kappa_2 = \sigma^2\kappa_1/m^2 + p_a p_b\{(\kappa_1)_a - (\kappa_1)_b\}^2. \tag{27}$$

https://doi.org/10.1017/S0305004100022994 Published online by Cambridge University Press

Part II

Application to sequential analysis

4. To apply the above theory to sequential analysis, we consider the random variable represented by $L_1 - L_0$, the logarithm of the likelihood ratio, as the size n of sample increases with time. Approximate values to take for a and b appropriate to specified risks of error have been inferred generally by Wald(5). With these values, the approximate formula for the expected size of sample is obtained from (20). In the large-sample theory case where the difference between the two values considered for an unknown parameter θ is small, the above theory becomes more self-contained, since the values to take for a and b may then be deduced from it. We write, when $\theta = \theta_1$,

$$L(\theta_1) - L(\theta_0) = (\theta_1 - \theta_0)\left(\frac{\partial L}{\partial \theta}\right)_1 - \tfrac{1}{2}(\theta_1 - \theta_0)^2 \left(\frac{\partial^2 L}{\partial \theta^2}\right)',$$

where $(\partial^2 L/\partial\theta^2)'$ denotes the value of $\partial^2 L/\partial\theta^2$ at θ' $(\theta_0 \leqslant \theta' \leqslant \theta_1)$, and as $\theta_0 - \theta_1$ tends to zero,

$$L(\theta_1) - L(\theta_0) \to (\theta_1 - \theta_0)\left(\frac{\partial L}{\partial \theta}\right)_1 + \frac{n}{2}(\theta_1 - \theta_0)^2 I(\theta), \tag{28}$$

where $I(\theta)$ is the value of Fisher's information function per observation in the neighbourhood of θ_1, θ_0. We have similarly when $\theta = \theta_0$,

$$L(\theta_1) - L(\theta_0) \to (\theta_1 - \theta_0)\left(\frac{\partial L}{\partial \theta}\right)_0 - \frac{n}{2}(\theta_1 - \theta_0)^2 I(\theta). \tag{29}$$

Hence (29) contains the reverse drift m to (28), where

$$2m = \sigma^2 = (\theta_1 - \theta_0)^2 I(\theta). \tag{30}$$

If, whether $\theta = \theta_1$ or θ_0, we cease sampling when $F \equiv L(\theta_1) - L(\theta_0)$ exceeds a, or is less than $-b$, and assert that $\theta = \theta_1$ in the former case and θ_0 in the latter, it follows from (25) that if the chance of error if $\theta = \theta_1$ is to be ϵ_1, and to be ϵ_2 if $\theta = \theta_0$, we must put

$$\alpha^2 = \tfrac{1}{2}\log\frac{1-\epsilon_1}{\epsilon_2}, \quad \beta^2 = \tfrac{1}{2}\log\frac{1-\epsilon_2}{\epsilon_1},$$

or since $\alpha^2 = am/\sigma^2$, $\beta^2 = bm/\sigma^2$, this corresponds to

$$F = \log\frac{1-\epsilon_1}{\epsilon_2} \quad \text{or} \quad \log\frac{\epsilon_1}{1-\epsilon_2}, \tag{31}$$

in the two cases, in agreement with Wald's theory.

We have immediately from (26),

$$\kappa_1(\theta_1 - \theta_0)^2 I(\theta) = 2\left\{(1-\epsilon_1)\log\frac{1-\epsilon_1}{\epsilon_2} + \epsilon_1 \log\frac{\epsilon_1}{1-\epsilon_2}\right\}. \tag{32}$$

Further large-sample formulae appropriate to values of the unknown parameter θ other than the basic values θ_1 and θ_0 are readily obtained by modifying the 'drift' by the factor

$$\frac{(\theta_1 - \theta)^2 - (\theta_0 - \theta)^2}{(\theta_1 - \theta_0)^2} \cdot \frac{\theta_1 + \theta_0 - 2\theta}{\theta_1 - \theta_0}.$$

https://doi.org/10.1017/S0305004100022994 Published online by Cambridge University Press

The case of more than one unknown

5. We shall consider for convenience the case of two unknowns θ and ϕ.

(i) If the rival hypotheses concern both θ and ϕ, we have at once

$$L(\theta_1,\phi_1)-L(\theta_0,\phi_0)=(\theta_1-\theta_0)\left(\frac{\partial L}{\partial\theta}\right)_1+(\phi_1-\phi_0)\left(\frac{\partial L}{\partial\phi}\right)_1$$
$$-\frac{1}{2}\left\{(\theta_1-\theta_0)^2\left(\frac{\partial^2 L}{\partial\theta^2}\right)'+2(\theta_1-\theta_0)(\phi_1-\phi_0)\left(\frac{\partial^2 L}{\partial\theta\partial\phi}\right)'+(\phi_1-\phi_0)^2\left(\frac{\partial^2 L}{\partial\phi^2}\right)'\right\}$$
$$\to X+nm, \tag{33}$$

where, if $\theta=\theta_1$, $\phi=\phi_1$, X is a random variable increasing in variance at the rate per observation of

$$2m=(\theta_1-\theta_0)^2 I_{\theta\theta}+2(\theta_1-\theta_0)(\phi_1-\phi_0)I_{\theta\phi}+(\phi_1-\phi_0)^2 I_{\phi\phi}, \tag{34}$$

where
$$I_{\theta\theta}=E\left(\frac{\partial L}{\partial\theta}\right)_1^2=E\left(-\frac{\partial^2 L}{\partial\theta^2}\right)_1, \quad \text{etc.}$$

(ii) If the rival hypotheses concern θ only, we may consider

$$F\equiv L(\theta_1,\hat{\phi}_1)-L(\theta_0,\hat{\phi}_0), \tag{35}$$

where $\hat{\phi}_1$ denotes the maximum likelihood estimate of ϕ given $\theta=\theta_1$, etc. For $\theta=\theta_1$ we have

$$F\to(\theta_1-\theta_0)\left(\frac{\partial L}{\partial\theta}\right)_1+(\hat{\phi}_1-\hat{\phi}_0)\left(\frac{\partial L}{\partial\phi}\right)_{\theta=\theta_1}$$
$$+\frac{n}{2}\{(\theta_1-\theta_0)^2 I_{\theta\theta}-2(\theta_1-\theta_0)(\hat{\phi}_0-\phi)I_{\theta\phi}+[(\hat{\phi}_0-\phi)^2-(\hat{\phi}_1-\phi)^2]I_{\phi\phi}\},$$

or since
$$\left(\frac{\partial L}{\partial\phi}\right)_0\to\left(\frac{\partial L}{\partial\phi}\right)_1+(\theta_1-\theta_0)I_{\theta\phi},$$

we have
$$\hat{\phi}_0\to\hat{\phi}_1+(\theta_1-\theta_0)I_{\theta\phi}/I_{\phi\phi}. \tag{36}$$

whence
$$F\to(\theta_1-\theta_0)\left\{\left(\frac{\partial L}{\partial\theta}\right)_1-I_{\theta\phi}\left(\frac{\partial L}{\partial\phi}\right)_{\theta=\theta_1}\Big/I_{\phi\phi}\right\}$$
$$+\frac{n}{2}\{(\theta_1-\theta_0)^2 I_{\theta\theta}-(\theta_1-\theta_0)^2 I_{\theta\phi}^2/I_{\phi\phi}\}$$
$$\to X+nm, \tag{37}$$

where, when $\theta=\theta_1$, X is a random variable increasing in variance at the rate per observation of $(\theta_1-\theta_0)^2 I(\theta\,|\,\phi)$, where

$$I(\theta\,|\,\phi)=I_{\theta\theta}-I_{\theta\phi}^2/I_{\phi\phi}. \tag{38}$$

I am indebted to the Chief Scientific Officer, Ministry of Supply, for permission to publish this paper.

REFERENCES

(1) CHANDRASEKHAR, S. *Rev. Mod. Phys.* 15 (1943), 1.
(2) FÜRTH, R. *Ann. d. Phys.* 53 (1917), 177.
(3) TWEEDIE, M. C. K. *Nature, London*, 155 (1945), 453.
(4) WALD, A. *Ann. Math. Statist.* 15 (1944), 283.
(5) WALD, A. *Ann. Math. Statist.* 16 (1945), 117.

QUEENS' COLLEGE
CAMBRIDGE

https://doi.org/10.1017/S0305004100022994 Published online by Cambridge University Press