Lecture V. HEURISTIC TREATMENT OF LARGE DEVIATIONS

I shall describe a simple but non-rigorous approach to the usual theory of large deviations for a sum of independent, identically distributed random variables when the moment-generating function exists in an interval about the origin. In this approach the moment-generating function enters by way of a linear approximation to the logarithm of the desired density function. This suggests the possible use of a quadratic approximation when a linear approximation is inadequate. It might be desirable to try to carry this suggestion a bit further, perhaps by numerical work in special cases. This may not be easy. Perhaps it is possible to make some of this work rigorous with the aid of the ideas of the sixth lecture, which is incomplete but presumably correct mathematically. My work on this lecture followed some comments by Frank Hampel at a lecture I gave in Zürich at the Eidgenössische Technische Hochschule. After some hesitation, I have decided to retain this lecture despite its very unsatisfactory state.

Let $X_1, X_2, \ldots$ be independently identically distributed real random variables with common probability density function p and, for all n, let p_n be the density function of

(1) $$W_n = \sum_{i=1}^{n} X_i .$$

Suppose

(2) $$EX_i = 0,$$

and let

(3) $$\alpha = -\log p,$$

(4) $$\alpha_n = -\log p_n,$$

(5) $$\phi(t) = E\, e^{tX_i} = \int \exp[tx - \alpha(x)]dx,$$

and

(6) $$\Psi = \log \phi.$$

It will be assumed that

(7) $$\phi(t) < \infty$$

for all $t \in (a,b)$ where a and b are real numbers such that

(8) $$a < 0 < b.$$

Then, for all n and w,

(9) $$\frac{w}{n+1} = E\left\{\frac{W_{n+1}}{n+1} \,\Big|\, W_{n+1} = w\right\} = \sum_{i=1}^{n+1} E\left\{\frac{W_i}{n+1} \,\Big|\, W_{n+1} = w\right\}$$

$$= E\{X_{n+1} | W_{n+1} = w\} = \frac{\int x p(x) p_n(w-x)dx}{\int p(x)p_n(w-x)dx}$$

$$= \frac{\int x \exp[-\alpha(x)-\alpha_n(w-x)]dx}{\int \exp[-\alpha(x)-\alpha_n(w-x)]dx}.$$

Proceeding non-rigorously, using the linear approximation

(10) $$\alpha_n(w-x) \approx \alpha_n(w) - \alpha_n'(w)x$$

we obtain

(11) $$\frac{w}{n+1} \approx \frac{\int x \exp[-\alpha(x)+\alpha_n'(w)x]dx}{\int \exp[-\alpha(x)+\alpha_n'(w)x]dx} = \Psi'(\alpha_n'(w)).$$

Using a local central limit theorem to conclude that

(12) $$\alpha_n(0) \approx \frac{1}{2}\log(2\pi\sigma^2 n),$$

where

(13) $$\sigma^2 = E\,X_i^2,$$

we obtain for $w > 0$

(14) $$\alpha_n(w) \approx \frac{1}{2}\log(2\pi\sigma^2 n) + \int_0^w \Psi'^{-1}\left(\frac{x}{n+1}\right)dx,$$

which gives the correct leading terms for α_n, the negative logarithm of the density of W_n.

Now let us carry the approximation (14) a bit further to obtain an expression for $\alpha_n(w)$ that is correct to within $O(\frac{1}{n})$ for $w = O(n)$ when the moment-generating function is finite in a neighborhood of the origin. Again the argument is not rigorous. Let us assume that

(15) $$\alpha_n'(w) = \rho_1\left(\frac{w}{n}\right) + \frac{1}{n}\rho_2\left(\frac{w}{n}\right) + O\left(\frac{1}{n^2}\right),$$

where ρ_1 and ρ_2 are smooth functions in the sense that (15) can be differentiated to obtain

(16) $$\alpha_n''(w) = \frac{1}{n}\rho_1'\left(\frac{w}{n}\right) + O\left(\frac{1}{n^2}\right).$$

Returning to (9), I shall use the more precise approximation

(17) $$\alpha_n(w-x) = \alpha_n(w) - \alpha_n'(w)x + \frac{1}{2}\alpha_n''(w)x^2 + O\left(\frac{1}{n^2}\right)$$

instead of (10). Then (9) yields

(18) $$\frac{w}{n+1} = \frac{\int x\exp[-\alpha(x) + x\alpha_n'(w) - \frac{1}{2}x^2\alpha_n''(w) + O(\frac{1}{n^2})]dx}{\int \exp[-\alpha(x) + x\alpha_n'(w) - \frac{1}{2}x^2\alpha_n''(w) + O(\frac{1}{n^2})]dx}$$

$$= \frac{\int x(1 - \frac{1}{2}x^2\alpha_n''(w) + O(\frac{1}{n^2}))\exp[-\alpha(x) + x\alpha_n'(w)]dx}{\int (1 - \frac{1}{2}x^2\alpha_n''(w) + O(\frac{1}{n^2}))\exp[-\alpha(x) + x\alpha_n'(w)]dx}$$

$$= \frac{\phi'(\alpha_n'(w)) - \frac{1}{2}\alpha_n''(w)\phi'''(\alpha_n'(w)) + O(\frac{1}{n^2})}{\phi(\alpha_n'(w)) - \frac{1}{2}\alpha_n''(w)\phi''(\alpha_n'(w)) + O(\frac{1}{n^2})}$$

$$= \Psi'(\alpha_n'(w)) - \alpha_n''(w)[\Psi'(\alpha_n'(w))\Psi''(\alpha_n'(w)) + \frac{1}{2}\Psi'''(\alpha_n'(w))] + O(\frac{1}{n^2}).$$

Solving this approximately, taking (15) and (16) into account, we obtain

$$\alpha_n'(w) = \Psi'^{-1}(\frac{w}{n+1} + \alpha_n''(w)(\Psi'\Psi'' + \frac{1}{2}\Psi''')(\alpha_n'(w))) + O(\frac{1}{n^2}) \tag{19}$$

$$= \Psi'^{-1}(\frac{w}{n+1}) + \frac{w}{(n+1)^2} \frac{1}{\Psi''(\Psi'^{-1}(\frac{w}{n+1}))} + \frac{1}{2(n+1)} \frac{\Psi'''}{\Psi''^2} (\Psi'^{-1}(\frac{w}{n+1})) + O(\frac{1}{n^2})$$

$$= \Psi'^{-1}(\frac{w}{n}) + \frac{1}{2n} \frac{\Psi'''}{\Psi''^2} (\Psi'^{-1}(\frac{w}{n})) + O(\frac{1}{n^2}),$$

leading to

$$\alpha_n(w) = C_n + \int_0^w \Psi'^{-1}(\frac{x}{n})dx + \frac{1}{2}\log \Psi''(\Psi'^{-1}(\frac{w}{n})) + O(\frac{1}{n}) \tag{20}$$

for $w = O(n)$, where C_n is an appropriately chosen constant. This agrees with the leading term in (for example) formula (26) of Daniels (1954) when we take into account the fact that

$$\frac{d}{dz}[\Psi(\Psi'^{-1}(z)) - z\Psi'^{-1}(z)] = -\Psi'^{-1}(z). \tag{21}$$

From his formula (or an appropriate form of the local central limit theorem) we also see that we can take $C_n = \frac{1}{2}\log(2\pi n)$.

Now let us look briefly at the possibility of using the local quadratic approximation (17) to α_n more carefully. It is required that $\alpha_n''(w)$ be positive. Let us introduce a function Ψ of two variables defined for real t and positive u by

$$\Psi(t,u) = \log E^{tX-uX^2} \tag{22}$$

$$= \log \int \exp(tx-ux^2-\alpha(x))dx.$$

Since the function $x \mapsto tx-ux^2$ is bounded above, this is finite for all real t and positive u. Thus, subject only to the condition that $\alpha_n''(w)$ is positive, we have, instead of (18)

(23)
$$\frac{w}{n+1} \approx \frac{\int x \exp[-\alpha(x)+x\alpha_n'(w) - \frac{1}{2} x^2\alpha_n''(w)]dx}{\int \exp[-\alpha(x)+x\alpha_n'(w) - \frac{1}{2} x^2\alpha_n''(w)]dx}$$

$$= \Psi_1(\alpha_n'(w),\alpha_n''(w)).$$

Thus, to a first approximation, we have a first-order differential equation for α_n'.