

A PROBABILITY LIMIT THEOREM REQUIRING NO MOMENTS

JOHN G. KEMENY¹

Let S_n be the sum of n independent, identically distributed, integer-valued random variables. Chung and Erdős² have shown that for arbitrary integers a, a', b ,

$$(1) \quad \lim_{n \rightarrow \infty} \frac{\Pr[S_{n-b} = a]}{\Pr[S_n = a']} = 1,$$

under one of two assumptions. They assume either that $E\{X\} = 0$, or that the mean diverges on both the positive and negative side. $E\{X\}$ denotes, as usual, the expectation—or mean—of the random variable X .

Actually their result is easily extendable to lattice-valued random variables, except that in the more general case a and a' must be restricted to be suitable values of the sums, and b must be so chosen and the limit taken over such a subsequence of n that the two probabilities be positive. One convenient general class to treat is the class of random variables whose values are multiples of a given rational number r , and for which S_n for sufficiently large n may be any multiple (or any nonnegative or any nonpositive multiple) of r . Then b may be any integer, a and a' multiples of the given rational number, and the limit is taken over sufficiently large n . We will prove a considerable extension of (1). For convenience we will restrict ourselves to rational valued random variables of the type indicated.

Let X take on the values a_i with probability p_i . We number the values in increasing order, indicating negative values with negative subscripts. All unrestricted sums will be over the entire range of X . The generating function $f(s) = \sum p_i s^{a_i}$ is well defined in some interval (s_1, s_2) , and may or may not converge at the end-points. We will restrict the domain to positive numbers, and hence we always have $0 \leq s_1 \leq 1 \leq s_2 \leq \infty$, though we may have the interval shrinking down to the single point 1.

We define $g(s) = sf'(s)/f(s)$. If $s_1 < s_2$, then this function is well

Presented to the Society, October 25, 1958; received by the editors November 3, 1958.

¹ The author wishes to acknowledge the aid of the National Science Foundation through a grant given to the Dartmouth Mathematics Project.

² *Probability limit theorems assuming only the first moment I*, by K. L. Chung and P. Erdős, *Memoirs Amer. Math. Soc.*, No. 6, 1951.

defined on the interior of the interval. Since $g(1) = \sum a_i p_i$, this may be well defined (possibly $\pm \infty$) even if $s_1 = s_2 = 1$. The only case in which g is not defined is where the mean diverges on both sides. A proof of our theorem for this case will be given separately. We will show that g is monotone increasing in the interior of the interval.

$$\begin{aligned} g'(s) &= [f'(s)f(s) + sf''(s)f(s) - sf'(s)^2]/f(s)^2 \\ &= \sum \sum [p_i p_j (a_i^2 - a_i a_j) s^{a_i + a_j - 1}]/f(s)^2 \\ &= 1/2 \sum \sum [p_i p_j (a_i - a_j)^2 s^{a_i + a_j - 1}]/f(s)^2 > 0. \end{aligned}$$

Let us introduce the quantities $t_1 = \lim_{s \rightarrow s_1^+} g(s)$ and $t_2 = \lim_{s \rightarrow s_2^-} g(s)$; t_1 may be $-\infty$ and t_2 may be $+\infty$. Then g maps (s_1, s_2) one-one onto (t_1, t_2) . We then define the function $h(t)$ as follows: If $t_1 < t < t_2$, $h(t) = g^{-1}(t)$. If $t \leq t_1$, $h(t) = t_1$, and if $t \geq t_2$, then $h(t) = t_2$. The function h is a mapping of all real numbers onto $[s_1, s_2]$. If $s_1 = s_2 = 1$, we let $h(t) = 1$ for all t .

In the following theorem, and from here on, we will always assume that a and a' are appropriate multiples of the rational number r , that t is a multiple of r between $\inf(a_i)$ and $\sup(a_i)$, and that limits are carried out for large enough values of n that the probabilities be positive. We do *not* assume the existence of any moments for X .

THEOREM.³ For permissible a, a', t ,

$$(2) \quad \lim_{n \rightarrow \infty} \frac{\Pr[S_{n-b} = nt + a]}{\Pr[S_n = nt + a']} = \frac{h(t)^{a'-a}}{E\{h(t)^X\}^b}.$$

PROOF. Let us first consider the case $t_1 < t < t_2$. In this case $h(t) = s_0$ is in the domain of f , and hence

$$E\{h(t)^X\} = f(h(t)) = c$$

is finite. Let us introduce the random variable X^* which takes on the value $a_i - t$ with probability $p_i^* = p_i s_0^{a_i} / c$. We have $p_i^* > 0$ and $\sum p_i^* = 1$ by the definition of c .

$$\begin{aligned} E\{X^*\} &= (1/c) \sum p_i s_0^{a_i} (a_i - t) \\ &= (1/c) [s_0 f'(s_0) - t f(s_0)] \\ &= (1/c) f(s_0) [g(s_0) - t]. \end{aligned}$$

But $s_0 = h(t)$ and hence $g(s_0) = t$. Thus $E\{X^*\} = 0$. Thus we may apply (1) to X^* , and we obtain

³ The existence of these limits was conjectured by J. L. Snell and R. E. Williamson. The author is indebted to them for suggesting the problem, and for other valuable suggestions. An analogous theorem for the case of a random variable with a continuous density function will be published by Snell and Williamson.

$$(3) \quad \lim_{n \rightarrow \infty} \frac{\Pr[S_{n-b}^* = a]}{\Pr[S_n^* = a']} = 1.$$

We obtain $\Pr[S_n^* = a']$ by choosing all paths in which the numbers n_i (the number of times $X^* = a_i - t$) satisfy the conditions $\sum n_i = n$ and $\sum n_i(a_i - t) = a'$. For each such path we have a probability $\prod p_i^{*n_i}$. Thus the desired probability is the sum of all these products over the chosen paths. But $\prod p_i^{*n_i} = (s_0^{nt+a}/c^n) \prod p_i^{n_i}$, where we made use of both properties of the chosen paths. We now observe that for each path $\sum n_i = n$ and $\sum n_i a_i = nt + a'$, hence each is a path on which $S_n = nt + a'$, and each path has probability $\prod p_i^{n_i}$ for X . Hence we obtain that

$$\Pr[S_n^* = a'] = (s_0^{nt+a'}/c^n) \Pr[S_n = nt + a']$$

and similarly

$$\Pr[S_{n-b}^* = a] = (s_0^{nt+a}/c^{n-b}) \Pr[S_{n-b} = nt + a].$$

Thus we may substitute into equation (3). We then note that the factors multiplying the probabilities simplify, and we obtain

$$\lim_{n \rightarrow \infty} (c^b/s_0^{a'-a}) \frac{\Pr[S_{n-b} = nt + a]}{\Pr[S_n = nt + a']} = 1.$$

From this we obtain our theorem by bringing the factor, which is free of n , to the right side, and replacing s_0 and c by the quantities they stand for.

The cases $t \leq t_1$ and $t \geq t_2$ are quite similar; we will treat only the latter. In this case we must proceed by a method of truncation. We introduce the random variables X_n which take on the values a_i for $i \leq n$ with probabilities $k_n p_i$. In other words, we cut off the high values of X , and renormalize the probabilities by a suitable factor k_n . The original X is the limit of the X_n , and we will show that the limit (2) may be obtained as the limit of the limits of the X_n .

Let us consider some X_n , with n sufficiently large so that t is permissible for it (that is $t < a_n$). Since X_n takes on only a finite number of positive values, f_n will have a domain which contains $(1, +\infty)$. Furthermore, $E\{X_n\}$ is always well defined, though it may be $-\infty$. As for any random variable having a well-defined mean, $g_n(1)$ is the mean (or approaches this mean as $s \rightarrow 1$ from above or below if the mean is $-\infty$ or $+\infty$ respectively). As $s \rightarrow \infty$, $g_n(s)$ tends to the largest value a_n of X_n . Hence the range of g_n includes all permissible values of t above the mean.

We will assume that $E\{X\}$ is well defined (though possibly in-

finite). The case where the mean diverges on both sides will be treated separately. Then $g(1)$ is the mean. Our $t \geq t_2 \geq g(1)$ is at least equal to the mean. Obviously, $E\{X\} > E\{X_n\}$, hence t is greater than the mean of X_n , and hence it lies in the range of g_n . Hence our previously proven result is applicable to X_n . For the ratio in (2), as applied to X_n , we obtain the desired limit in terms of $h_n(t)$.

We will show below that $h_n(t) \rightarrow h(t) = s_2$. We would like to take this limiting h -value as the value in (2). For this we must estimate the error between the ratio in (2) and $s_2^{a'-a}/E\{s_2^X\}^b$. We can bind this error by the sum of three terms, each of which can be shown to go to 0: (i) The difference between the ratio for X and for some X_N . This will go to 0 for $N(n)$ chosen sufficiently great. (ii) The difference between the ratio for X_N and its limit. We need a bound on this that is uniform in N . This may be obtained by a computation similar to that in the Chung-Erdős paper. (iii) The difference between $h_n(t)$ and s_2 , which tends to 0 since the sequence converges to s_2 .

We must still prove that $h_n(t) \rightarrow s_2$. Since t is in the range of each g_n , we have $g_n(h_n(t)) = t$. We note that

$$g_n(s) = \frac{\sum_{-\infty}^n k_n p_i a_i s^{a_i}}{\sum_{-\infty}^n k_n p_i s^{a_i}}.$$

In this ratio the normalizing factor cancels out. If $s > s_2$ is outside the domain of g , then both numerator and denominator will tend to ∞ . Hence the positive tails will dominate. But in the tail the numerator has terms arbitrarily larger than the corresponding terms in the denominator, hence $g_n(s) \rightarrow \infty$. Let s be any number greater than s_2 . Then for sufficiently large n we will have $g_n(s) > t$, and hence $h_n(t) < s$. Thus for any $\epsilon > 0$, all sufficiently high n yield $h_n(t) < s_2 + \epsilon$. But we also know that $g(s_2) = t_2 < t$. Thus $\sum p_i (a_i - t) s^{a_i} < 0$. If $a_n > t$, then the corresponding sum for X_n has some positive terms deleted and hence is certainly negative. Hence $g_n(s_2) < t$ for sufficiently high n . Hence $h_n(t) > s_2$ for sufficiently high n . But then $h_n(t) \rightarrow s_2$, as was to be shown.

It is not immediately obvious that $E\{h(t)^X\} = f(s_2)$ is finite in this case. Indeed, in general $f(s_2)$ is not finite. However, in the present case we know that $g(s_2) = t_2$ is finite. If $f(s) \rightarrow \infty$ as $s \rightarrow s_2$, then we can use the same argument that we used for g_n above to show that $g(s) \rightarrow \infty$, contrary to our assumption.

This leaves us only one case, where the mean diverges on both sides. In this case $s_1 = s_2 = 1$, and hence $h(t) = 1$ for all t . We must

show that in this case we always obtain the limit 1 in (2). But this is a trivial consequence of (1). Let $E\{X\}$ diverge on both sides. Then the same is true for $X-t$. If we apply (1) to $X-t$, using $a+bt$ in place of a , we obtain (2).

This completes the proof of our theorem.

It is worth pointing out that the above result is a genuine generalization of the Chung-Erdős result. If the mean is not defined, then we have $h(t)=1$ for all t , and hence (1) is a special case of (2) in which $t=0$ is chosen. If the mean is some finite number t_0 , then $g(1)=t_0$, hence $h(t_0)=1$, and we obtain limit 1 for $t=t_0$. The mean 0 case in (1) is a special case of this fact.

Let us illustrate the theorem by means of some examples. If X takes on only a finite number of values, then f has domain $(-\infty, +\infty)$. The range of g over $(0, \infty)$ will be (t_1, t_2) where t_1 is the smallest and t_2 the largest value of X . Hence $h(t)=g^{-1}(t)$ for all permissible t . Hence $h(t)$ is strictly monotone increasing in t . This means that for $b=0$, $a=0$, $a'=1$ the limit $h(t)$ is strictly increasing; while for $b=1$, $a=a'=0$ the limit $E\{h(t)^x\}^{-1}$ is strictly decreasing. For example, for coin tossing we have $a_0=0$, $a_1=1$, $p_0=q$, $p_1=p$, $f(s)=q+ps$, $g(s)=ps/(q+ps)$, $h(t)=qt/p(1-t)$, $E\{h(t)^x\}=q/(1-t)$. The limit in (2) is $q^{a'-a-b}t^{a'-a}/p^{a'-a}(1-t)^{a'-a-b}$.

An example of infinitely many values, where all permissible t are in the range of g , is given by the Poisson distribution: For $0 \leq s < \infty$, $0 \leq t < \infty$,

$$f(s) = \sum_{i=0}^{\infty} e^{-\lambda} \frac{\lambda^i}{i!} s^i = e^{\lambda(s-1)}, \quad g(s) = s\lambda,$$

$$h(t) = t/\lambda, \quad E\{h(t)^x\} = e^{t-\lambda}.$$

An example where we have permissible t -values outside the range of g is the following:

$$f(s) = \frac{1}{\zeta(3)} \sum_{i=1}^{\infty} \frac{1}{i^3} s^i, \quad 0 \leq s \leq 1,$$

$$g(s) = \frac{\sum_{i=1}^{\infty} \frac{1}{i^2} s^i}{\sum_{i=1}^{\infty} \frac{1}{i^3} s^i},$$

$$t_2 = g(1) = \zeta(2)/\zeta(3).$$

Hence for $t > \zeta(2)/\zeta(3)$, $h(t)=1$. For all such t we obtain limits 1 in (2). Then for example,

$$\frac{\Pr[S_n = nt]}{\Pr[S_n = nt + 1]}$$

increases monotonely for $0 < t \leq \zeta(2)/\zeta(3)$ and is 1 from then on.

There is an interesting alternative way to interpret the process determined by X^* . Let us compute the conditional probability that the first outcome was a_i , given that the sum is nt . Denote this probability by \bar{p}_i^n . Then

$$\bar{p}_i^n = \Pr[X = a_i | S_n = nt].$$

That is

$$\bar{p}_i^n = \frac{\Pr[X = a_i] \Pr[S_{n-1} = nt - a_i]}{\Pr[S_n = nt]}.$$

By (2) we have

$$\lim_{n \rightarrow \infty} \bar{p}_i^n = \frac{p_i h(t)^{a_i}}{E\{h(t)^X\}}.$$

If $t_1 < t < t_2$, then this limiting distribution is the same as the distribution of $X^* + t$. Hence it has mean t . If t is outside this range, for example $t \geq t_2$, then the limiting distribution is equal to the distribution obtained for $t = t_2$. If t equals the mean of the original distribution, then $h(t) = 1$ and hence $\bar{p}_i = p_i$.

For any distribution, the above procedure produces a one parameter family of distributions. This family includes the original distribution for t equal to its mean. It is easy to verify that this family of distributions is closed in the sense that if any one element is chosen for the starting distribution the same family is obtained.⁴

As an example of this let us take the Poisson distribution with mean λ as our starting distribution. Then

$$\bar{p}_i^n = \Pr[X = i | S_n = nt] = \binom{nt}{i} (1/n)^i (n - 1/n)^{nt-i}.$$

That is, we obtain the binomial measure for nt trials with probability $1/n$ for each trial. As n tends to infinity this distribution approaches the Poisson distribution with mean $\lambda = t$. Note that we can obtain this classical limiting theorem directly from our results. If we choose $t = \lambda$ we know that we obtain the starting process. That is we obtain the Poisson measure with $\lambda = t$. Hence the above binomial measure for $t = \lambda$ approaches the Poisson measure when n tends to infinity.

⁴ This family of distributions was suggested to the author by J. L. Snell.