UCRL 549
Cy. 2

UNIVERSITY OF CALIFORNIA

Radiation Laboratory

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 5545

BERKELEY, CALIFORNIA

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Copy No. 2

UNCLASSIFIED

UCRL-549

UNIVERSITY OF CALIFORNIA

Radiation Laboratory

Contract No. W-7405-eng-48

THE DISTRIBUTION OF THE POSITIVE-NEGATIVE RATIO FOR MESONS

Bayard Rankin

December 14, 1949

Berkeley, California

The Distribution of the Positive-Negative Ratio for Mesons

Bayard Rankin

December 14, 1949

Positive and negative mesons produced in a common target in the cyclotron are received in two photographic plates. In the magnetic field of the cyclotron, negatives are deflected to one plate, positives to the other. The mesons must pass through an absorber before entering the plates in order to be observable. The experimental design is that used by Dr. Hugh Bradner for determining the positive-negative ratio for mesons. Observations are made on the plates by counting the number of mesons and taking account of the area covered.

The problem is (1) to estimate the true ratio of positive to negative mesons produced in nuclear reactions under given conditions; (2) to find the statistical error in the estimate. No assumptions need be made about the processes involved in the nuclear reactions.

The problem is solved by finding the maximum likelihood estimate of the ratio and by deriving the distribution of a variable related to the ratio of two Poissons. In the second part of the discussion the contents of a new theorem is encompassed and the theorem is proved.

Let n be total number of mesons produced in one bombardment.

Let p be the probability that a positive meson is produced, and (1 - p) the probability that a negative is produced.

Let p' be the probability that a meson is observed (in a unit square of the plate) once it has been produced. For reasoning purposes, say we confine our observations to one unit square on each emulsion plate.

Assumption (1)*: p' is the same for all mesons, and is constant.

Let the production of a meson be called a trial, so that there are n trials in a bombardment.

We shall say:

$$P_I = Pr\left\{ \text{seeing (+) meson in one trial} \right\} = pp'$$

$$P_{II} = Pr\left\{ \text{seeing (-) meson in one trial} \right\} = (1 - p)\, p'$$

$$P_{III} = Pr\left\{ \text{seeing nothing} \right\} = 1 - p'$$

For n trials, let ξ be the number of (+) mesons observed

η the number of (-).

Then

$$Pr\left\{\xi = \ell, \eta = k\right\} = \frac{n!}{\ell!k!(n-\ell-k)!} P_I^{\ell} P_{II}^{k} P_{III}^{n-\ell-k} = P_{\ell k}$$

Write:

$$P_I = \frac{nP_I}{n} = \frac{np'p}{n} = \frac{\lambda p}{n} \quad \text{where } \lambda = np'$$

$$P_{II} = \frac{nP_{II}}{n} = \frac{np'(1-p)}{n} = \frac{\lambda(1-p)}{n}$$

$$P_{III} = 1 - p' = 1 - \frac{\lambda}{n}$$

Assumption (2): In one bombardment $n \longrightarrow \infty$

Assumption (3): p' is very small, i.e., $\lambda = np' = \text{const.}$

We want to find $P_{\ell k}$ under assumptions (2) and (3). To do this, take limits in such a way that $\lambda = np'$ is constant as $n \longrightarrow \infty$. The procedure

---

*Note: At least two things are included under this assumption. First, we assume that positive and negative mesons do not interact with matter differently while passing through the absorber. Second, we assume that the observer is no more apt to overlook one type of particle than another, whether it is a positive, a negative star-former, or a negative non-star-former. Actually, a separate correction must be made on the ideal solution, if, as is customary, the observer looks only for π-μ decays and negative stars. Positives decay approximately 95% of the cases; negatives form stars in approximately 73% of the cases.

is as follows:

$$P_{\ell k} = \frac{n!}{\ell!k!(n-\ell-k)!}\left(\frac{\lambda p}{n}\right)^{\ell}\left(\frac{\lambda(1-p)}{n}\right)^{k}\left(1-\frac{\lambda}{n}\right)^{n-\ell-k}$$

$$= \frac{n(n-1)\ldots(n-\ell-k+1)}{\ell!k!}\left(\frac{\lambda p}{n}\right)^{\ell}\left(\frac{\lambda(1-p)}{n}\right)^{k}\left(1-\frac{\lambda}{n}\right)^{n-\ell-k}$$

$$= \left(\frac{(\lambda p)}{\ell!}\right)^{\ell}\left(\frac{\{\lambda(1-p)\}^{k}}{k!}\right)\left[\frac{(1-1/n)\ldots(1-\frac{\ell+k-1}{n})}{\left(1-\frac{\lambda}{n}\right)^{\ell+k}}\right]\left[1-\frac{\lambda}{n}\right]^{n}$$

In the limit as $n \longrightarrow \infty$

$$P_{\ell k} \longrightarrow \frac{(\lambda p)^{\ell}}{\ell!}\left(\frac{\{\lambda(1-p)\}^{k}}{k!}\right)e^{-\lambda}$$

Consequently, though ξ and η were dependent variables previously with joint probability law $P_{\ell k}$ (n finite), they become independent in the limit with joint probability law:

$$\Pr\left\{\xi = \ell, \eta = k\right\} = \frac{(\lambda p)^{\ell}}{\ell!}e^{-\lambda p} \cdot \frac{\{\lambda(1-p)\}^{k}}{k!}e^{-\lambda(1-p)}$$

Each variable follows a Poisson distribution.

In order to estimate $R = \frac{p}{1-p}$ we will have to consider the distribution of the ratio of ξ and η. Two difficulties are involved here. If we want to find the distribution directly, we must deal with summations rather than integrals; moreover, the denominator of our new variable will have a finite probability of becoming zero.

To obviate these difficulties, we might consider continuous random variables directly related with ξ and η. The same problem, that of estimating R, will be treated. The mathematics will be simpler and only slight modifications need be made on the procedure of observation. Suppose we look for mesons only over an area where p' is constant. Instead of observing the

number of particles per unit area, observe the area covered in finding each particle.

Since the total number of particles situated in a given area will be distributed as a Poisson, let λp be the average number of (+) particles per unit area. Let X be a random variable, which is the distance traveled from an arbitrary point to the point of discovery of the first (+) particle. Unit breadth is covered in traversing distance.

Then X will be distributed according to the following law:

$$\Pr\left\{ X \leq x \right\} = \int_0^x \lambda p e^{-\lambda p x}\, dx = 1 - e^{-\lambda p x}$$

A similar definition for Y for negative particles will give:

$$\Pr\left\{ Y \leq y \right\} = \int_0^y \lambda(1 - p)\, e^{-\lambda(1 - p)y}\, dy$$

The characteristic function for X and Y will be the same except for the parameters involved. That for X is:

$$\phi_x(T) = E(itx) = \int_0^\infty \lambda p e^{x(it - \lambda p)}\, dx = \lambda p \int_0^\infty \left[e^{-\lambda p x} \cos tx + i e^{-\lambda p x} \sin tx \right] dx$$

$$= \frac{1}{1 + \frac{t^2}{(\lambda p)^2}} + i \frac{\frac{t}{(\lambda p)}}{1 + \left(\frac{t}{\lambda p}\right)^2} = \frac{1}{1 - \frac{it}{\lambda p}}$$

The characteristic function for the mean will be:

$$\phi_{\bar{x}}(t) = \left[\phi_x \left(\frac{t}{n}\right) \right]^n = \left[\frac{1}{1 - \frac{it}{n \lambda p}} \right]^n$$

This is the characteristic function for a variable following the Γ-distribution.

Hence, we know that the distribution for $\bar{X}$ and $\bar{Y}$ are:

$$\Pr\left\{ \bar{X} \leq x \right\} = \int_0^x \frac{(n \lambda p)^n\, x^{n-1}\, e^{-n \lambda p x}}{\Gamma(n)}\, dx$$

$$\Pr\left\{\overline{Y} \leq y\right\} = \int_0^y \frac{[m\lambda(1-p)]^m y^{m-1} e^{-m\lambda(1-p)y}}{\Gamma(m)} dy$$

And their joint distribution will be:

$$\Pr\left\{\overline{X} \leq x, \overline{Y} \leq y\right\} = \int_0^x \int_0^y \frac{(n\lambda p)^n [m\lambda(1-p)]^m}{\Gamma(n)\Gamma(m)} x^{n-1} y^{m-1} e^{-n\lambda px - m\lambda(1-p)y} dxdy$$

$$= \int_0^x \int_0^y f(x,y)\, dxdy$$

It can be shown by maximizing the joint density function of the variables $(X_1, X_2, \ldots X_n, Y_1, Y_2, \ldots Y_m)$ that the maximum likelihood estimate of the parameter p (which is the probability of producing a positive meson) is given by:

$$\hat{p} = \frac{\frac{1}{m}\sum_{i=1}^{m} y_i}{\frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{m}\sum_{i=1}^{m} y_i}$$

More simply, we can show that we get the same result by maximizing f(x,y), where x, y are the means.

$$(1)\begin{cases} \dfrac{\partial \log f}{\partial p} = \dfrac{n}{p} - \dfrac{m}{(1-p)} - n\lambda x - m\lambda y \\[2ex] \dfrac{\partial \log f}{\partial \lambda} = \dfrac{n}{\lambda} + \dfrac{m}{\lambda} - npx - m(1-p)y \end{cases}$$

Setting equations (1) equal to zero and solving for $\hat{p}$, the maximizing value, we get:

$$p = \frac{y}{x+y} \qquad \frac{\hat{p}}{1-p} = \frac{n\sum y_i}{m\sum x_i} = \frac{y}{x} = u$$

u may or may not be an unbiased estimate of R and we cannot yet present it as a solution to the first part of our problem. It may be noticed that as the number of observations becomes greater the variable u approaches

the statistic

$$R' = \frac{n/N}{m/M},$$

where numerator and denominator are the usual estimates for the true number of particles per unit area based on N and M observations, or, in other words, the usual estimates based on continuous observations extending over areas of size N and M. Moreover, though the y_i's, and x_i's are the distances from arbitrary points to points of discovery of particles, they can also be distances from particle to particle, since the position of one discovery does not influence that the next. Our y's and x's will remain independent even though we make a continuous sweep across the plate and our procedure of scanning can be exactly as if we were to use the more intuitive R'. The one modification in the calculation of the sample ratio which distinguishes u from R' is the following modification (1):

> The areas $\sum y_i, \sum x_i$ must be the areas covered from an arbitrary starting point to the point of discovery of the last particle and no more.

Of course, this modification will be insignificant as more mesons are counted, but for small samples we will generally have $u < R'$ when $m < n$. We must keep in mind, however, that the foregoing calculations are made under the assumption that the particles are randomly distributed over the plate and the distances between them are directly related to one Poisson distribution for the x's, one for the y's. It may be that this assumption is superfluous, if equal areas on positive and negative plates are scanned.

The next step in our problem is to find the distribution of

$$\frac{\hat{p}}{1 - \hat{p}} = \frac{y}{x} = u$$

To do this make the transformation T on the variables following f(x,y) and find the distribution of u by integrating out the other variable.

$$f(x,y) = \frac{(n\lambda p)^n [m\lambda(1-p)]^m}{\Gamma(n)\Gamma(m)} x^{n-1} y^{m-1} e^{-n\lambda px - m\lambda(1-p)y}$$

$$T : \begin{cases} y = uv \\ x = v \end{cases}, \quad T^{-1} : \begin{cases} u = \frac{y}{x} \\ v = x \end{cases}, \quad |J| = v$$

$$f(u,v) = \frac{(n\lambda p)^n [m\lambda(1-p)]^m}{\Gamma(n)\Gamma(m)} u^{m-1} v^{n+m-1} e^{-v\{n\lambda p + m\lambda(1-p)n\}}$$

$$Z = v\left\{n\lambda p - m\lambda(1-p)u\right\}$$

$$v = \frac{Z}{n\lambda p - m\lambda(1-p)u}$$

$$f(u) = \frac{(n\lambda p)^n [m\lambda(1-p)]^m u^{m-1}}{\Gamma(n)\Gamma(m)[n\lambda p + m\lambda(1-p)u]^{n+m}} \int_0^\infty Z^{n+m-1} e^{-Z} dZ$$

$$= \frac{\Gamma(n+m)}{\Gamma(n)\Gamma(m)} \frac{(np)^n [m(1-p)]^m u^{m-1}}{[np + m(1-p)u]^{n+m}} = \frac{\left(\frac{m}{nR}\right)^m u^{m-1}}{\beta(n,m)\left[1 + \frac{mu}{nR}\right]^{n+m}}$$

This distribution has not been used before, but observe the distribution of the variable F where:

$$u = FR$$

$$f(F) = \frac{\left(\frac{m}{n}\right)^m F^{m-1}}{\beta(n,m)\left[1 + \frac{m}{n}F\right]^{n+m}} = \frac{m^m n^n F^{m-1}}{\beta(n,m)\left[n + mF\right]^{n+m}}$$

$F = \frac{u}{R}$ follows the F-distribution with 2m, 2n degrees of freedom.

We now see that a second modification will bring an answer to part (1), for

$$E(u) = \left(\frac{n}{n-1}\right)R .$$

Modification (2): Use the statistic

$$u' = \left(\frac{n-1}{n}\right) u = \frac{(n-1)\sum y_i}{m\sum x_i}$$

for which

$$E(u') = R \, .$$

Again, as in modification (1), the effect is slight for large samples but in general gives:

$$u' < u < R' \, .$$

Very fortunately, the distribution, f(u), is independent of λ, and the only parameter R can be treated as a scale parameter. Consequently, we can easily make statements about probable error, and can construct confidence intervals. For instance,

Let

$$F_2 \text{ be such that } Pr\left\{\frac{u}{R} < F_2\right\} = .95$$

$$F_1 \text{ be such that } Pr\left\{\frac{u}{R} > F_1\right\} = .95 \, ,$$

where we refer to the F-tables to establish these inequalities, then:

$$Pr\left\{F_1 < \frac{u}{R} < F_2\right\} = Pr\left\{\frac{u}{F_2} < R < \frac{u}{F_1}\right\} = .90$$

This is a 90 percent confidence interval for R. That is, in 90 percent of the cases, intervals constructed from the sample in this way will cover the true ratio R.

By known theorems, it can be deduced that:

$$t = \frac{\ell n\,(u/R) + (1/2m - 1/2n)}{\sqrt{2\,(1/2m + 1/2n)}}$$

approaches the Gaussian variable, with mean 0, variance 1, as $n, m \longrightarrow \infty$

Hence, for large m, n, say greater than 100, we can affirm:

$$Pr\left\{-.6745 < \frac{\ell n \frac{u}{R}}{\sqrt{2\,(1/2m + 1/2n}} < .6745\right\} \cong .50$$

i.e.,

$$\Pr\left\{ ue^{-.6745\sqrt{(1/m + 1/n)}} < R < ue^{.6745\sqrt{1/m + 1/n}} \right\} \cong .50$$

When m,n are large and m differs much from n, greater accuracy is achieved by writing:

$$\Pr\left\{ ue^{-.6745\sqrt{1/m + 1/n} + (1/2m - 1/2n)} < R < ue^{.6745\sqrt{1/m + 1/n} + (1/2m - 1/2n)} \right\} \cong .50$$

This gives a simple method of approximating statistical error when m,n are large.

In the course of this discussion we have discovered and proved a special case of the following general theorem.

<u>Theorem</u>: Let $A_1 = \sum y_i$ be the area covered in finding m particles distributed according to a Poisson distribution of parameter λ_1 (λ_1 being the average number of particles per unit area). Let $A_2 = \sum x_i$ have a corresponding definition for n particles of Poisson distribution and parameter λ_2. Then

$$F = \left(\frac{\lambda 1}{\lambda^2}\right) u, \text{ where } u = \frac{1/m \sum_{i=1}^{m} y_i}{1/n \sum_{i=1}^{n} x_i}$$

is a random variable which follows Snedecor's F-distribution with 2m, 2n degrees of freedom.

With m,n becoming increasingly large, the variable

$$t = 1/2\, \ln\left(\frac{\lambda_1 u}{\lambda_2}\right)$$

approaches the Gaussian variate with variance:

$$\sigma^2 = 1/2\ (1/2m + 1/2n) \qquad \text{and mean } -1/2\ (1/2m - 1/2n)$$

When the experiment is more refined and theoretical considerations single out one definite R_0 as the true ratio, we may wish to test the statistical hypothesis that $R = R_0$. The results obtained in this paper may be used as a first step in constructing the test.

The same statistical analysis may be applied to meson positive-negative ratio studies which do not make use of magnetic separation. It may also be used in connection with other problems where a ratio of production of two types of particles is under consideration, and where it can be established that

$$\frac{\lambda_2}{\lambda_1} = R.$$

Information Division
ead/12-15-49