The table below compares the exact formula (3) with the approximate formula (4) for some typical values of n and p. The ratio takes the same value for $p = x$ and $p = 1-x$, so values of p are restricted to the range $0{\cdot}1-(0{\cdot}1)-0{\cdot}5$. It will be noted that the Normal limiting value of $\sqrt{(2/\pi)} = 0{\cdot}7979$ is approached quite rapidly as n increases. The approximate formula (4) gives nearly three figure accuracy for $0{\cdot}2 \leqslant p \leqslant 0{\cdot}8$, while for $n \geqslant 50$ the exact and approximate values agree to four decimal places.

n \ p	0·1		0·2		0·3		0·4		0·5	
	(3)	(4)	(3)	(4)	(3)	(4)	(3)	(4)	(3)	(4)
10	0·7351	0·7313	0·7640	0·7630	0·7733	0·7729	0·7771	0·7768	0·7782	0·7779
20	0·7652	0·7642	0·7807	0·7804	0·7855	0·7854	0·7874	0·7874	0·7879	0·7879
50	0·7846		0·7909		0·7929		0·7937		0·7939	
100	0·7912		0·7949		0·7954		0·7958		0·7959	

REFERENCE

Gruder, O. (1930). *9th International Congress of Actuaries*, vol. II, p. 222.

A comment on D. V. Lindley's statistical paradox

By M. S. BARTLETT

University of Manchester

I read with considerable interest the discussion by D. V. Lindley (1957), demonstrating the possibility of contradiction between the result of a statistical significance test and an assessment of the posterior probability of a null hypothesis. I would agree that he establishes the point that one must be cautious when using a fixed significance level for testing a null hypothesis irrespective of the size of sample one is taking. However, there is a slip, in his expression for K under his equation (1), that appears to me, unless corrected, to lead to an overstatement of this point. The prior distribution for θ, given that $\theta \neq \theta_0$, was assumed to be uniform over an interval I, and hence its density function should be $1/I$ in this interval. This leads to the extra factor $1/I$ in the second term in the expression for K.* This expression then becomes consistent with Jeffreys's equation (10), § 5·0, in his book (second edition, 1948).

The occurrence of I in the formula for the posterior probability $\bar{c}$ of the null hypothesis θ_0, this quantity $\bar{c}$ satisfying approximately the relation

$$\frac{\bar{c}}{1-\bar{c}} = \frac{c}{1-c}\left[\frac{I}{\sigma}\sqrt{\left(\frac{n}{2\pi}\right)}\exp\left\{\frac{-n(\bar{x}-\theta_0)^2}{2\sigma^2}\right\}\right], \tag{1}$$

where c is the prior probability of θ_0, now makes the value of $\bar{c}$ much more arbitrary. In fact, in situations where one might be tempted to put I infinity the silly answer $\bar{c} = 1$ ensues. D. V. Lindley has suggested to me, in correspondence, that one way out of this dilemma would be to make $c/(1-c)$ the prior odds in favour of the null hypothesis against *any unit interval* of the alternative values, but this is rather an artificial evasion of the difficulty. It is common for those who use the Bayes's approach to assume a uniform prior distribution for the mean of a normal population. If the difference in means between two populations (for simplicity, of known equal variances) is considered, the question might legitimately be asked whether these populations, from each of which a sample is available, are identical. The most natural prior probabilities would seem to me, if we try to use the Bayes's approach, to be c for this null hypothesis, and a remaining uniform prior distribution of the true difference in population means over the entire infinite range.

The other point that might be noticed about formula (1), if we disregard the above difficulty and agree to leave I finite, is this. Certainly, for a fixed significance level (and I and σ fixed), the posterior odds increase with $\sqrt{n}$. But from the Neyman-Pearson theory of the power of tests, if a null hypothesis θ_0 is being tested against a single alternative θ_1, the sample size n would if possible be chosen in relation to the 'distance' $d \equiv \theta_1 - \theta_0$ between the two hypotheses ($\sqrt{n}$ inversely proportional to d). The situation under discussion is a little more complicated but, with a range of I for the alternatives, it would be fairly

Downloaded from http://biomet.oxfordjournals.org/ at University of Chicago on April 30, 2014

reasonable to choose the sample size n analogously, making $\sqrt{n}$ proportional to $1/I$. If we write $\sqrt{n} = A\sigma/I$, we obtain

$$\frac{\bar{c}}{1-\bar{c}} = \frac{c}{1-c}\left\{\frac{A}{\sqrt{(2\pi)}}e^{-\frac{1}{2}\lambda^2}\right\}, \tag{2}$$

where $\lambda = \sqrt{n}(\bar{x}-\theta_0)/\sigma$; in (2) there is a constant relation between $\bar{c}$ and λ for fixed c.

REFERENCES

Jeffreys, H. (1948). *Theory of Probability*, 2nd ed. Oxford: Clarendon Press.
Lindley, D. V. (1957). A statistical paradox. *Biometrika*, **44**, 187–92.

Editorial Note. Mr Lindley agrees and apologizes for the fact that a factor $1/I$ was omitted from his equation (1), but points out that in the two examples which he discusses this factor is unity. His general argument as to the limiting value of $\bar{c}$ is in any case unaffected, and his two particular examples are also unaffected. There appears to be no real difference of opinion between Prof. Bartlett and Mr Lindley on this point.

The point raised by Prof. Bartlett's second paragraph is related to the difficulty of laying down a uniform prior probability for a parameter of infinite range, a point which, in my opinion, has not been properly cleared up; if the probability of μ is $kd\mu$, integration over the infinite range leads to the conclusion that k is zero (the integral having to be unity). The root of this difficulty seems to be that several limiting processes are involved and no clear rules have been laid down as to which, if any, has priority. In any case, this point mainly concerns estimation, whereas Mr Lindley was concerned with testing hypotheses.

In regard to Prof. Bartlett's final point, it may be useful to observe that some procedure of the type he suggests is implicit in the idea of the asymptotic relative efficiency of a test. In general, the power of a test against a specific alternative tends to unity with increasing sample size. To compare two tests asymptotically at a fixed significance level, it is necessary to allow the alternative to approach the null hypothesis as the sample size increases.

M.G.K.

* There is also a further dropping of a factor $1/\sigma$ in the last formula on p. 191, but this is a more trivial slip.

CORRIGENDA

Biometrika (1957), **44**, pp. 150–8

'The use of a concomitant variable in collecting an experimental design.' By D. R. Cox

Dr K. R. Nair has kindly pointed out that some of the results for Methods II and V of the above paper have been given by him in *Sankhya* (1942), **6**, 167–174. He has also noted that in formula (6) of my paper $(\bar{x}_i-\bar{x}_j)^2$ should read $k(\bar{x}_i-\bar{x}_j)^2$.

D.R.C.

Biometrika (1957), **44**, pp. 168–78

'Multiple runs.' By D. E. Barton and F. N. David

P. 174, line 10. *Delete* 'transition probabilities' and *substitute* 'joint probabilities of two successive events'.

P. 174, line 13. *Delete* the full stop after the expression for W and *add* 'and where $p_i = 1/k$, $(i = 1, 2, \ldots, k)$'.

Downloaded from http://biomet.oxfordjournals.org/ at University of Chicago on April 30, 2014