DOCUMENT RESUME

ED 095 224 TM 003 911

AUTHOR Lord, Frederic M.
TITLE A Statistical Test for Cheating.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RM-74-7
PUB DATE Apr 74
NOTE 9p.

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Cheating; *Hypothesis Testing; *Statistical Analysis; *Testing; *Tests of Significance; True Scores

ABSTRACT

A statistical test for cheating is developed. The case of a single examinee who has taken parallel forms of the same selection test on three occasions, obtaining scores x, y, z, is used to illustrate the development. It is assumed that each score is normally distributed with the same known variance, that is, the variance of the errors of measurement. These scores are further assumed to be distributed independently, since each score differs from its mean (true) value only because of errors of measurement. Based on these assumptions, a significance test is presented to indicate evidence of cheating. Mathematical derivations for the test of significance are presented as well as a numerical example. (Author/RC)

RM-74-7

RESEARCH MEMORANDUM

A STATISTICAL TEST FOR CHEATING

Frederic M. Lord

This Memorandum is for interoffice use. It is not to be cited as a published report without the specific permission of the author.

Educational Testing Service
Princeton, New Jersey
April 1974

A Statistical Test for Cheating

1. Introduction

A single examinee has taken parallel forms of the same selection test on three occasions, obtaining scores x , y , z . We assume that each score is normally distributed with the same known variance σ^2 , the variance of the errors of measurement. The scores are assumed to be distributed independently, since each score differs from its mean ('true') value only because of an error of measurement.

Either

H_0 : all three scores have the same mean μ

or

H_1 : scores x and z have mean μ and score y has mean ν ($\nu > \mu$).

We wish a significance test for the null hypothesis H_0 . In practice, rejection of H_0 is considered evidence of cheating on test y .

2. Reparameterization

Define $\theta \equiv \nu - \mu$ and consider θ and μ as the (unknown) parameters. Now H_0 is $\theta = 0$. This is in standard form for a composite hypothesis about θ . The unspecified ('nuisance') parameter is μ .

3. Transformation of Sample Space

Let us replace the sample observations x , y , z by the transformed observations

(1) $$m \equiv (x + y + z)/3 \quad ,$$

(2) $$t \equiv (y - m)/\sigma\sqrt{2/3}$$

$$\equiv (-x + 2y - z)/\sigma\sqrt{6} \quad ,$$

(3) $$d = (y - x)/\sigma\sqrt{2} \quad .$$

This transformation will turn out to be useful in setting up the desired significance test. (The reason x and z are not treated symmetrically in (3) is discussed in Section 10.)

4. Distribution of Sample under H_0

From (1), (2), and (3) we find the means, variances, and covariances of m, t, and d under H_0 to be

(4) $$\begin{cases} \mu_{om} = \mu \quad , & \sigma_m^2 = \sigma^2/3 \quad , \\ \mu_{ot} = 0 \quad , & \sigma_t^2 = 1 \quad , \\ \mu_{od} = 0 \quad , & \sigma_d^2 = 1 \quad , \\ \sigma_{dt} = \frac{1}{2}\sqrt{3} \quad , & \sigma_{mt} = \sigma_{md} = 0 \quad . \end{cases}$$

The joint distribution of m, t, and d under H_0 is normal trivariate with the parameters given by (4).

5. A Sufficient Statistic under H_0

Since t and d have zero covariance with m under H_0, they are distributed independently of m. It appears from (4) that under H_0, the joint distribution of t and d does not depend on μ . Thus m is a sufficient statistic for μ under H_0 .

6. Uniformly Most Powerful Significance Test

Because of the sufficiency of m for the nuisance parameter μ , the problem of finding a uniformly most powerful significance test for the composite hypothesis $H_0 : \Theta = 0$ is reduced to the problem of finding a best critical region for testing the simple hypothesis that $\Theta = 0$, m being held constant (Kendall and Stuart, 1973, section 17.20; Lehmann, 1959, section 4.3). The best critical region for the simple hypothesis will depend on the conditional distributions of the transformed observations for given m .

7. Conditional Distributions of Observations

Both t and d have zero covariance with m under either H_0 or H_1 . Thus the distribution of t and d is not affected by holding m constant. Under H_1 , the means of t and d are seen to be

$$(5) \qquad \begin{cases} \mu_{1t} = \Theta/\sigma\sqrt{3/2} \, , \\ \mu_{1d} = \Theta/\sigma\sqrt{2} \, . \end{cases}$$

The variances and covariances under H_1 are the same as those under H_0 , listed in (4).

Under H_1 , the joint distribution $f_1(t,d)$ of t and d (whether conditional on m or unconditional) is normal bivariate so that

$$(6) \qquad \log f_1(t,d) = K - \frac{1}{2(1-\rho^2)}[(t-\mu_{1t})^2 + (d-\mu_{1d})^2$$

$$- 2\rho(t-\mu_{1t})(d-\mu_{1d})]$$

where K is a constant and

$$\rho = \sigma_{dt}/\sigma_d\sigma_t = \frac{1}{2}\sqrt{3} \quad .$$

Under H_0, μ_{1t} and μ_{1d} are replaced by $\mu_{ot} = \mu_{od} = 0$, so the corresponding equation under H_0 is

$$(7) \qquad \log f_o(t,d) = K - \frac{1}{2(1-\rho^2)}(t^2 + d^2 - 2\rho td) \quad .$$

8. Likelihood Ratio

From (6) and (7) the logarithm of the likelihood ratio is

$$\ell \equiv \log f_o(t,d) - \log f_1(t,d)$$

$$\ell \propto -2t\mu_{1t} + \mu_{1t}^2 - 2d\mu_{1d} + \mu_{1d}^2$$

$$+ 2\rho t\mu_{1d} + 2\rho d\mu_{1t} - 2\rho\mu_{1d}\mu_{1t} \quad .$$

Substituting from (4) and (5),

$$\ell \propto -2t\theta/\sigma\sqrt{3/2} + 2\theta^2/3\sigma^2 - 2d\theta/\sigma\sqrt{2} + \theta^2/2\sigma^2$$

$$+ \sqrt{3}\, t\theta/\sigma\sqrt{2} + \sqrt{3}\, d\theta/\sigma\sqrt{3/2} - \sqrt{3}\,\theta^2/\sigma^2\sqrt{3} \quad .$$

The terms involving d drop out (this was the reason for choosing definition (2) for t) and after simplification

$$(8) \qquad \ell \propto \theta/\sigma\sqrt{6} - t \quad .$$

9. Best Critical Region

The best critical region for the simple hypothesis, and thus the uniformly most powerful significance test for the composite hypothesis, is defined by $\ell <$ constant when H_0 holds (Kendall and Stuart, 1973, section 22.10; Lehmann, 1959, section 3.3). Since $\theta = 0$ under H_0, the best critical region can be defined by $t \geq k_0$, the constant k_0 being chosen so that $\text{Prob}(t \geq k_0 | H_0) = \alpha$, the chosen significance level.

10. Truncation on d

In practical work, it is sometimes a practice to make no investigation of an examinee unless $d > d_0$ where d_0 is some predetermined value (this is the reason for dealing with the rather awkward variable d throughout this report). This avoids searching out z for large numbers of individuals for whom it is highly unlikely that $t \geq k_0$.

Under this truncation $d > d_o$, the situation in section 4 is the same as before except that now d and t have a singly truncated normal bivariate distribution, independent of m, the truncation being on d. Since d has zero-covariance with m, restriction on d does not prevent m from being a sufficient statistic for μ under H_0. Thus, the problem is still reduced to the problem of finding the best critical region for the same simple hypothesis.

The bivariate distribution of t and d is now proportional to $f_1(t,d)$ of (6) except that f_1 is zero when $d < d_o$. A similar change occurs for $f_o(t,d)$, so that the likelihood ratio λ remains unaltered. The critical region is still defined by $t \geq k_o$, where now

$$\mathrm{Prob}(t \geq k_o | H_0, \ d \geq d_o) = \alpha$$

implies either a different k_o for given α or a different α for given k_o than was used in the previous section.

11. Small Sample Properties

It should be noted that the best critical region and the uniformly most powerful significance test are chosen for their optimum properties, which do not require large sample size. It can be shown that t is a sufficient statistic for Θ when m is fixed.

12. Numerical Example

It is known from reliability studies that the errors of measurement of Scholastic Aptitude Test scores have a standard deviation of roughly $\sigma = 13\sqrt{6} \doteq 32$ on the CEEB score scale. Suppose $x = 400$, $y = 610$, $z = 430$. Now $m = 480$, $t = 5$. Under H_0, t has a mean of 0 and a standard deviation of 1. Thus $\text{Prob}(t \geq 5 | H_0)$ is .0000003.

References

Kendall, M. G. and Stuart, A. The advanced theory of statistics. Vol. 2 (3rd ed.). New York: Hafner, 1973.

Lehmann, E. L. Testing statistical hypotheses. New York: Wiley, 1959.