Lady tasting tea

In the design of experiments in statistics, the **lady tasting tea** is a randomized experiment devised by Ronald Fisher and reported in his book *The Design of Experiments* (1935).[1] The experiment is the original exposition of Fisher's notion of a null hypothesis, which is "never proved or established, but is possibly disproved, in the course of experimentation".[2][3]

The lady in question (Muriel Bristol) claimed to be able to tell whether the tea or the milk was added first to a cup. Fisher proposed to give her eight cups, four of each variety, in random order. One could then ask what the probability was for her getting the specific number of cups she identified correct, but just by chance.

Fisher's description is less than 10 pages in length and is notable for its simplicity and completeness regarding terminology, calculations and design of the experiment.[4] The example is loosely based on an event in Fisher's life. The test used was Fisher's exact test.

The experiment asked whether a taster could tell if the milk was added before the brewed tea, when preparing a cup of tea

The experiment

The experiment provided the lady with 8 randomly ordered cups of tea—4 prepared by first adding milk, 4 prepared by first adding the tea. She was to select the 4 cups prepared by one method. This offered the lady the advantage of judging cups by comparison. She was fully informed of the experimental method.

The null hypothesis was that the lady had no ability to distinguish the teas. In Fisher's approach, there is no alternative hypothesis,[2] unlike in the Neyman–Pearson approach.

Ronald Fisher in 1913

The test statistic was a simple count of the number of successes in selecting the 4 cups (the number of cups of the given type successfully selected). The distribution of possible numbers of successes, assuming the null hypothesis was true, was computed using the number of permutations. Using the combination formula, with $n = 8$ total cups and $k = 4$ cups chosen, there are $\frac{8!}{4!(8-4)!} = 70$ possible combinations.

Tea-Tasting Distribution Assuming the Null Hypothesis

Success count	Permutations of selection	Number of permutations
0	oooo	1 × 1 = 1
1	ooox, ooxo, oxoo, xooo	4 × 4 = 16
2	ooxx, oxox, oxxo, xoxo, xxoo, xoox	6 × 6 = 36
3	oxxx, xoxx, xxox, xxxo	4 × 4 = 16
4	xxxx	1 × 1 = 1
Total		70

The frequencies of the possible numbers of successes, given in the final column of this table, are derived as follows. For 0 successes, there is clearly only one set of four choices (namely, choosing all four incorrect cups) giving this result. For one success and three failures, there are four correct cups of which one is selected, which by the combination formula can occur in $\binom{4}{1} = 4$ different ways (as shown in column 2, with x denoting a correct cup that is chosen and o denoting a correct cup that is not chosen); and independently of that, there are four incorrect cups of which three are selected, which can occur in $\binom{4}{3} = 4$ ways (as shown in the second column, this time with x interpreted as an incorrect cup which is not chosen, and o indicating an incorrect cup which is chosen). Thus a selection of any one correct cup and any three incorrect cups can occur in any of 4×4 = 16 ways. The frequencies of the other possible numbers of successes are calculated correspondingly. Thus the number of successes is distributed according to the hypergeometric distribution.

The critical region for rejection of the null of no ability to distinguish was the single case of 4 successes of 4 possible, based on the conventional probability criterion < 5%. This is the critical region because under the null of no ability to distinguish, 4 successes has 1 chance out of 70 (≈ 1.4% < 5%) of occurring, whereas at least 3 of 4 successes has a probability of (16+1)/70 (≈ 24.3% > 5%).

Thus, if and only if the lady properly categorized all 8 cups was Fisher willing to reject the null hypothesis – effectively acknowledging the lady's ability at a 1.4% significance level (but without quantifying her ability). Fisher later discussed the benefits of more trials and repeated tests.

David Salsburg reports that a colleague of Fisher, H. Fairfield Smith, revealed that in the test, the woman got all eight cups correct.[5][6] The chance of someone who just guesses getting all correct, assuming she guesses that four had the tea put in first and four the milk, would be only 1 in 70 (the combinations of 8 taken 4 at a time).

In popular science, Salsburg published a book entitled *The Lady Tasting Tea*,[5] which describes Fisher's experiment and ideas on randomization. Deb Basu wrote that "the famous case of the 'lady tasting tea'" was "one of the two supporting pillars ... of the randomization analysis of experimental data."[7]

See also

- Hypergeometric distribution
- Permutation test
- Random assignment
- Randomization test

References

1. Fisher 1971, II. The Principles of Experimentation, Illustrated by a Psycho-physical Experiment.
2. Fisher 1971, Chapter II. The Principles of Experimentation, Illustrated by a Psycho-physical Experiment, Section 8. The Null Hypothesis.
3. OED quote: **1935** R. A. Fisher, *The Design of Experiments* ii. 19, "We may speak of this hypothesis as the 'null hypothesis', and it should be noted that the null hypothesis is never proved or established, but is possibly disproved, in the course of experimentation."
4. Fisher, Sir Ronald A. (1956) [*The Design of Experiments* (1935)]. "Mathematics of a Lady Tasting Tea". In James Roy Newman. *The World of Mathematics, volume 3* (https://books.google.com/?id=oKZwtLQTmNAC&pg=PA1512&dq=%22mathematics+of+a+lady+tasting+tea%22). Courier Dover Publications. ISBN 978-0-486-41151-4.
5. Salsburg (2002)
6. Box, Joan Fisher (1978). *R.A. Fisher, The Life of a Scientist*. New York: Wiley. p. 134. ISBN 0-471-09300-9.
7. Basu (1980a, p. 575; 1980b)

- Fisher, Ronald A. (1971) [1935]. *The Design of Experiments* (9th ed.). Macmillan. ISBN 0-02-844690-9.
- Basu, D. (1980a). "Randomization Analysis of Experimental Data: The Fisher Randomization Test". *Journal of the American Statistical Association*. **75** (371): 575–582. doi:10.2307/2287648 (https://doi.org/10.2307%2F2287648). JSTOR 2287648 (https://www.jstor.org/stable/2287648).
- Basu, D. (1980b). "The Fisher Randomization Test", reprinted with a new preface in *Statistical Information and Likelihood : A Collection of Critical Essays by Dr. D. Basu* ; J. K. Ghosh, editor. Springer 1988.
- Kempthorne, Oscar. "Intervention experiments, randomization and inference". In Malay Ghosh and Pramod K. Pathak. *Current Issues in Statistical Inference – Essays in Honor of D. Basu* (http://projecteuclid.org/euclid.lnms/1215458836). Hayward, CA.: IMS. pp. 13–31. doi:10.1214/lnms/1215458836 (https://doi.org/10.1214%2Flnms%2F1215458836).
- Salsburg, D. (2002) *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*, W.H. Freeman / Owl Book. ISBN 0-8050-7134-2