

Hosmer–Lemeshow test

From Wikipedia, the free encyclopedia

The **Hosmer–Lemeshow test** is a statistical test for goodness of fit for logistic regression models. It is used frequently in risk prediction models. The test assesses whether or not the observed event rates match expected event rates in subgroups of the model population. The Hosmer–Lemeshow test specifically identifies subgroups as the deciles of fitted risk values. Models for which expected and observed event rates in subgroups are similar are called well calibrated.

The Hosmer–Lemeshow test statistic is given by:

$$H = \sum_{g=1}^G \frac{(O_g - E_g)^2}{N_g \pi_g (1 - \pi_g)}.$$

Here O_g , E_g , N_g , and π_g denote the observed events, expected events, observations, predicted risk for the g^{th} risk decile group, and G is the number of groups. The test statistic asymptotically follows a χ^2 distribution with $G - 2$ degrees of freedom. The number of risk groups may be adjusted depending on how many fitted risks are determined by the model. This helps to avoid singular decile groups.

References

- Hosmer, David W.; Lemeshow, Stanley (2000). *Applied Logistic Regression*. New York: Wiley. ISBN 0-471-61553-6.
- Agresti, Alan (2002). *Categorical Data Analysis*. Hoboken: John Wiley and Sons. ISBN 0-471-36093-7.

Retrieved from "https://en.wikipedia.org/w/index.php?title=Hosmer–Lemeshow_test&oldid=599927753"

Categories: Statistical tests | Regression analysis

-
- This page was last modified on 16 March 2014, at 22:46.
 - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.