

Wald test

In statistics, the **Wald test** (named after Abraham Wald) assesses constraints on statistical parameters based on the weighted distance between the unrestricted estimate and its hypothesized value under the null hypothesis, where the weight is the precision of the estimate.^{[1][2]} Intuitively, the larger this weighted distance, the less likely it is that the constraint is true. While the finite sample distributions of Wald tests are generally unknown,^[3] it has an asymptotic χ^2 -distribution under the null hypothesis, a fact that can be used to determine statistical significance.^[4]

Together with the Lagrange multiplier and the likelihood-ratio test, the Wald test is one of three classical approaches to hypothesis testing. An advantage of the Wald test over the other two is that it only requires the estimation of the unrestricted model, which lowers the computational burden as compared to the likelihood-ratio test. However, a major disadvantage is that (in finite samples) it is not invariant to changes in the representation of the null hypothesis; in other words, algebraically equivalent expressions of non-linear parameter restriction can lead to different values of the test statistic.^{[5][6]} That is because the Wald statistic is derived from a Taylor expansion,^[7] and different ways of writing equivalent nonlinear expressions lead to nontrivial differences in the corresponding Taylor coefficients.^[8] Another aberration, known as the Hauck–Donner effect, can occur in binomial models when the estimated (unconstrained) parameter is close to the boundary of the parameter space—for instance a fitted probability being extremely close to zero or one—which results in the Wald test no longer monotonically increasing in the distance between the unconstrained and constraint parameter.^[9]

Contents

Mathematical details
Test on a single parameter
Test(s) on multiple parameters
Nonlinear hypothesis
Non-invariance to re-parameterisations
Alternatives to the Wald test
See also
References
Further reading
External links

Mathematical details

Under the Wald test, the estimated $\hat{\theta}$ that was found as the maximizing argument of the unconstrained likelihood function is compared with a hypothesized value θ_0 . In particular, the squared difference $\hat{\theta} - \theta_0$ is weighted by the curvature of the log-likelihood function.

Test on a single parameter

If the hypothesis involves only a single parameter restriction, then the Wald statistic takes the following form:

$$W = \frac{(\hat{\theta} - \theta_0)^2}{\text{var}(\hat{\theta})}$$

which under the null hypothesis follows an asymptotic χ^2 -distribution with one degree of freedom. The square root of the single-restriction Wald statistic can be understood as a (pseudo) t -ratio that is, however, not actually t -distributed except for the special case of linear regression.^[10] Instead, it follows an asymptotic z distribution.^[11]

$$\sqrt{W} = \frac{\hat{\theta} - \theta_0}{\text{se}(\hat{\theta})}$$

where $\text{se}(\hat{\theta})$ is the standard error of the maximum likelihood estimate (MLE), the square root of the variance. There are several ways to consistently estimate the variance matrix which in finite samples leads to alternative estimates of standard errors and associated test statistics and p -values.^[12]

Test(s) on multiple parameters

The Wald test can be used to test a single hypothesis on multiple parameters, as well as to test jointly multiple hypotheses on single/multiple parameters. Let $\hat{\theta}_n$ be our sample estimator of P parameters (i.e., $\hat{\theta}_n$ is a $P \times 1$ vector), which is supposed to follow asymptotically a normal distribution with covariance matrix V , $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, V)$. The test of Q hypotheses on the P parameters is expressed with a $Q \times P$ matrix R :

$$\begin{aligned} H_0 : R\theta &= r \\ H_1 : R\theta &\neq r \end{aligned}$$

The test statistic is:

$$(R\hat{\theta}_n - r)' [R(\hat{V}_n/n)R']^{-1} (R\hat{\theta}_n - r) \xrightarrow{D} \chi_Q^2$$

where \hat{V}_n is an estimator of the covariance matrix.^[13]

Proof

Suppose $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, V)$. Then, by Slutsky's theorem and by the properties of the normal distribution, multiplying by R has distribution:

$$R\sqrt{n}(\hat{\theta}_n - \theta) = \sqrt{n}(R\hat{\theta}_n - r) \xrightarrow{D} N(0, RVR')$$

Recalling that a quadratic form of normal distribution has a Chi-squared distribution:

$$\sqrt{n}(R\hat{\theta}_n - r)' [RVR']^{-1} \sqrt{n}(R\hat{\theta}_n - r) \xrightarrow{D} \chi_Q^2$$

Rearranging n finally gives:

$$(R\hat{\theta}_n - r)' [R(V/n)R']^{-1} (R\hat{\theta}_n - r) \xrightarrow{D} \chi_Q^2$$

What if the covariance matrix is not known a-priori and needs to be estimated from the data? If we have a consistent estimator $\hat{V}_n \sim \mathbf{X}_{n-P}^2$ of V , then by the independence of the covariance estimator and equation above, we have:

$$(R\hat{\theta}_n - r)' [R(\hat{V}_n/n)R']^{-1} (R\hat{\theta}_n - r) \xrightarrow{D} F(Q, n - P)$$

Nonlinear hypothesis

In the standard form, the Wald test is used to test linear hypotheses that can be represented by a single matrix R . If one wishes to test a non-linear hypothesis of the form:

$$\begin{aligned} H_0 &: c(\theta) = 0 \\ H_1 &: c(\theta) \neq 0 \end{aligned}$$

The test statistic becomes:

$$c(\hat{\theta}_n)' \left[c'(\hat{\theta}_n) (\hat{V}_n/n) c'(\hat{\theta}_n)' \right]^{-1} c(\hat{\theta}_n) \xrightarrow{D} \chi_Q^2$$

where $c'(\hat{\theta}_n)$ is the derivative of c evaluated at the sample estimator. This result is obtained using the delta method, which uses a first order approximation of the variance.

Non-invariance to re-parameterisations

The fact that one uses an approximation of the variance has the drawback that the Wald statistic is not-invariant to a non-linear transformation/reparametrisation of the hypothesis: it can give different answers to the same question, depending on how the question is phrased.^{[14][5]} For example, asking whether $R = 1$ is the same as asking whether $\log R = 0$; but the Wald statistic for $R = 1$ is not the same as the Wald statistic for $\log R = 0$ (because there is in general no neat relationship between the standard errors of R and $\log R$, so it needs to be approximated).^[15]

Alternatives to the Wald test

There exist several alternatives to the Wald test, namely the likelihood-ratio test and the Lagrange multiplier test (also known as the score test). Robert F. Engle showed that these three tests, the Wald test, the likelihood-ratio test and the Lagrange multiplier test are asymptotically equivalent.^[16] Although they are asymptotically equivalent, in finite samples, they could disagree enough to lead to different conclusions.

There are several reasons to prefer the likelihood ratio test or the Lagrange multiplier to the Wald test:^{[17][18][19]}

- Non-invariance: As argued above, the Wald test is not invariant to a reparametrization, while the Likelihood ratio tests will give exactly the same answer whether we work with R , $\log R$ or any other monotonic transformation of R .^[5]
- The other reason is that the Wald test uses two approximations (that we know the standard error, and that the distribution is χ^2), whereas the likelihood ratio test uses one approximation (that the distribution is χ^2).
- The Wald test requires an estimate under the alternative hypothesis, corresponding to the "full" model. In some cases, the model is simpler under the zero hypothesis, so that one might prefer to use the score test (also called Lagrange Multiplier test), which has the advantage that it can be formulated in situations where the variability is difficult to estimate; e.g. the Cochran–Mantel–Haenszel test is a score test.^[20]

See also

- Chow test
- Sequential probability ratio test
- Sup-Wald test
- Student's t-test
- Welch's *t*-test

References

1. Fahrmeir, Ludwig; Kneib, Thomas; Lang, Stefan; Marx, Brian (2013). *Regression : Models, Methods and Applications*. Berlin: Springer. p. 663. ISBN 978-3-642-34332-2.

2. Ward, Michael D.; Ahlquist, John S. (2018). *Maximum Likelihood for Social Science : Strategies for Analysis* (<https://books.google.com/books?id=iqRyDwAAQBAJ&pg=PA36>). Cambridge University Press. p. 36. ISBN 978-1-316-63682-4.

3. Martin, Vance; Hurn, Stan; Harris, David (2013). *Econometric Modelling with Time Series: Specification, Estimation and Testing* (https://books.google.com/books?id=I_W4UCYJhrAC&pg=PA138). Cambridge University Press. p. 138. ISBN 978-0-521-13981-6.
4. Davidson, Russell; MacKinnon, James G. (1993). "The Method of Maximum Likelihood : Fundamental Concepts and Notation". *Estimation and Inference in Econometrics*. New York: Oxford University Press. p. 89. ISBN 0-19-506011-3.
5. Gregory, Allan W.; Veall, Michael R. (1985). "Formulating Wald Tests of Nonlinear Restrictions" (<https://lib.uwo.ca/cgi/viewcontent.cgi?article=1791&context=economicsresrpt>). *Econometrica*. **53** (6): 1465–1468. JSTOR 1913221 (<https://www.jstor.org/stable/1913221>).
6. Phillips, P. C. B.; Park, Joon Y. (1988). "On the Formulation of Wald Tests of Nonlinear Restrictions". *Econometrica*. **56** (5): 1065–1083. JSTOR 1911359 (<https://www.jstor.org/stable/1911359>).
7. Hayashi, Fumio (2000). *Econometrics* (<https://books.google.com/books?id=QyIW8WUlyzcC&pg=PA489>). Princeton: Princeton University Press. pp. 489–491. ISBN 1-4008-2383-8.
8. Lafontaine, Francine; White, Kenneth J. (1986). "Obtaining Any Wald Statistic You Want". *Economics Letters*. **21** (1): 35–40. doi:10.1016/0165-1765(86)90117-5 (<https://doi.org/10.1016%2F0165-1765%2886%2990117-5>).
9. Hauck, Walter W., Jr.; Donner, Allan (1977). "Wald's Test as Applied to Hypotheses in Logit Analysis". *Journal of the American Statistical Association*. **72** (360a): 851–853. doi:10.1080/01621459.1977.10479969 (<https://doi.org/10.1080%2F01621459.1977.10479969>).
10. Cameron, A. Colin; Trivedi, Pravin K. (2005). *Microeconometrics : Methods and Applications* (<https://books.google.com/books?id=Zf0gCwxC9ocC&pg=PA137>). New York: Cambridge University Press. p. 137. ISBN 0-521-84805-9.
11. Davidson, Russell; MacKinnon, James G. (1993). "The Method of Maximum Likelihood : Fundamental Concepts and Notation". *Estimation and Inference in Econometrics*. New York: Oxford University Press. p. 89. ISBN 0-19-506011-3.
12. Martin, Vance; Hurn, Stan; Harris, David (2013). *Econometric Modelling with Time Series : Specification, Estimation and Testing* (https://books.google.com/books?id=I_W4UCYJhrAC&pg=PA129). New York: Cambridge University Press. p. 129. ISBN 978-0-521-13981-6.
13. Harrell, Frank E., Jr. (2001). "Section 9.3.1". *Regression modeling strategies*. New York: Springer-Verlag. ISBN 0387952322.
14. Fears, Thomas R.; Benichou, Jacques; Gail, Mitchell H. (1996). "A reminder of the fallibility of the Wald statistic". *The American Statistician*. **50** (3): 226–227. doi:10.1080/00031305.1996.10474384 (<https://doi.org/10.1080%2F00031305.1996.10474384>).
15. Critchley, Frank; Marriott, Paul; Salmon, Mark (1996). "On the Differential Geometry of the Wald Test with Nonlinear Restrictions". *Econometrica*. **64** (5): 1213–1222. JSTOR 2171963 (<https://www.jstor.org/stable/2171963>).
16. Engle, Robert F. (1983). "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics". In Intriligator, M. D.; Griliches, Z. (eds.). *Handbook of Econometrics*. II. Elsevier. pp. 796–801. ISBN 978-0-444-86185-6.
17. Harrell, Frank E., Jr. (2001). "Section 9.3.3". *Regression modeling strategies*. New York: Springer-Verlag. ISBN 0387952322.
18. Collett, David (1994). *Modelling Survival Data in Medical Research*. London: Chapman & Hall. ISBN 0412448807.
19. Pawitan, Yudi (2001). *In All Likelihood*. New York: Oxford University Press. ISBN 0198507658.
20. Agresti, Alan (2002). *Categorical Data Analysis* (2nd ed.). Wiley. p. 232. ISBN 0471360937.

Further reading

- Greene, William H. (2012). *Econometric Analysis* (Seventh international ed.). Boston: Pearson. pp. 155–161. ISBN 978-0-273-75356-8.
- Kmenta, Jan (1986). *Elements of Econometrics* (<https://archive.org/details/elementsofeconom0003kmen>) (Second ed.). New York: Macmillan. pp. 492–493. ISBN 0-02-365070-2.
- Thomas, R. L. (1993). *Introductory Econometrics: Theory and Application* (Second ed.). London: Longman. pp. 73–77. ISBN 0-582-07378-2.

External links

- Wald test (<http://jeff560.tripod.com/w.html>) on the Earliest known uses of some of the words of mathematics (<http://jeff560.tripod.com/mathword.html>)

Retrieved from "https://en.wikipedia.org/w/index.php?title=Wald_test&oldid=916844112"

This page was last edited on 20 September 2019, at 22:51 (UTC).

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.