

MODEL CONSISTENCY TESTING*

Don Johnson

This work is produced by The Connexions Project and licensed under the Creative Commons Attribution License †

In many situations, we seek to check consistency of the observations with some preconceived model. Alternative models are usually difficult to describe parametrically since inconsistency may be beyond our modeling capabilities. We need a test that accepts consistency of observations with a model or rejects the model without pronouncing a more favored alternative. Assuming we know (or presume to know) the probability distribution of the observations under \mathcal{M}_0 , the models are

- $\mathcal{M}_0: r \sim p_{r|\mathcal{M}_0}(r)$
- $\mathcal{M}_1: r \approx p_{r|\mathcal{M}_0}(r)$

Null hypothesis testing seeks to determine if the observations are consistent with this description. The best procedure for consistency testing amounts to determining whether the observations lie in a highly probable region as defined by the null probability distribution. However, no one region defines a probability that is less than unity. We must restrict the size of the region so that it best represents those observations maximally consistent with the model while satisfying a performance criterion. Letting P_F be a false-alarm probability established by us, we define the decision region \mathfrak{R}_0 to satisfy

$$Pr[r \in \mathfrak{R}_0 | \mathcal{M}_0] = \int p_{r|\mathcal{M}_0}(r) dr = 1 - P_F$$

and

$$\min_{\mathfrak{R}_0} \left\{ \mathfrak{R}_0, \int \mathfrak{R}_0 dr \right\}$$

Usually, this region is located about the mean, but may not be symmetrically centered if the probability density is skewed. Our null hypothesis test for model consistency becomes

$$r \in \mathfrak{R}_0 \Rightarrow \text{"say observations are consistent"}$$

$$r \notin \mathfrak{R}_0 \Rightarrow \text{"say observations are not consistent"}$$

Example 1

Consider the problem of determining whether the sequence $r_l, l \in \{1, \dots, L\}$, is white and Gaussian with zero mean and unit variance. Stated this way, the alternative model is not provided: is this model correct or not? We could estimate the probability density function of the observations and test the estimate for consistency. Here we take the null-hypothesis testing approach of converting this problem into a one-dimensional one by considering the statistic $r = \sum_l r_l^2$, which has a χ_L^2 .

*Version 1.5: Aug 11, 2003 3:55 pm GMT-5

†<http://creativecommons.org/licenses/by/1.0>

Because this probability distribution is unimodal, the decision region can be safely assumed to be an interval $[r', r'']$.¹ In this case, we can find an analytic solution to the problem of determining the decision region. Letting $R = r'' - r'$ denote the width of the interval, we seek the solution of the constrained optimization problem

$$\min_{r', R} \{r', R\} \quad \text{subject to} \quad P_r(r' + R) - P_r(r') = 1 - P_F$$

We convert the constrained problem into an unconstrained one using Lagrange multipliers.

$$\min_{r', R} \{r', R + \lambda(P_r(r' + R) - P_r(r') - (1 - P_F))\}$$

Evaluation of the derivative of this quantity with respect to r' yields the result $p_r(r' + R) = p_r(r')$: to minimize the interval's width, the probability density function's values at the interval's endpoints must be equal. Finding these endpoints to satisfy the constraints amounts to searching the probability distribution at such points for increasing values of R until the required probability is contained within. For $L = 100$ and $P_F = 0.05$, the optimal decision region for the χ_L^2 distribution is $[78.82, 128.5]$. Figure 1 demonstrates ten testing trials for observations that fit the model and for observations that don't.

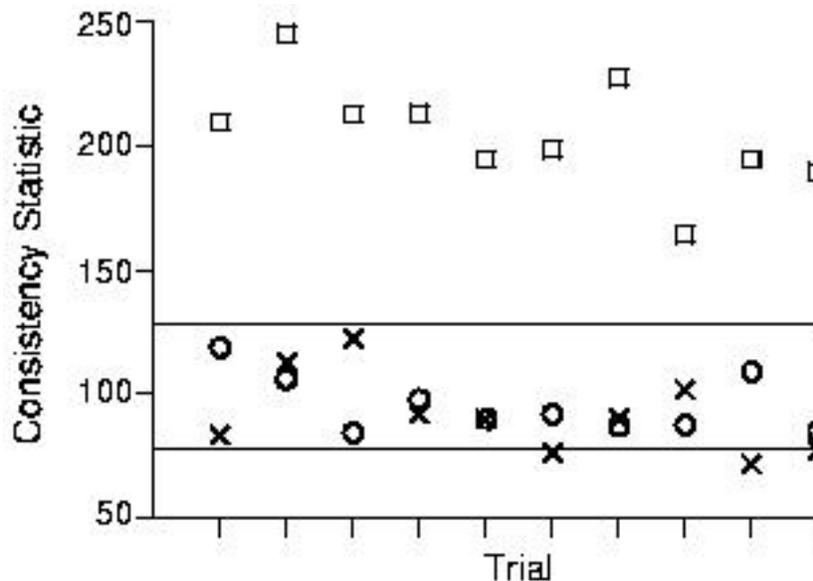


Figure 1: Ten trials of testing a 100-element sequence for consistency with a white, Gaussian model - $r_l \sim \mathcal{N}(0, 1)$ - for three situations. In the first (shown by the circles), the observations do conform to the model. In the second (boxes), the observations are zero-mean Gaussian but with variance two. Finally, the third example (crosses) has white observations with a density closely resembling the Gaussian: a hyperbolic secant density having zero mean and unit variance. The sum of squared observations for each example are shown with the optimal χ_{100}^2 interval displayed. Note how dramatically the test statistic departs from the decision interval when parameters disagree.

¹This one-dimensional result for the consistency test may extend to the multi-dimensional case in the obvious way.