Tests for Genetic Linkage and Homogeneity

Julian Faraway; Mohamed Lemdani; Odile Pons

Biometrics is currently published by International Biometric Society.

Dr. Li questions the appropriateness of our application of Bartlett's test of homogeneity and feels the correct test should be one on the equality of means rather than covariances. We used the test in the standard manner, rejected homogeneity, and therefore presented separate analyses of the genes with large numbers of codons, described the differences, and stopped. The additional test of equality of means would have been necessary if the hypothesis on quality of the covariances matrices had not been rejected.

George E. Booney
Department of Biostatistics
7701 Burholme Avenue
Philadelphia, Pennsylvania 19111, U.S.A.

Tests for Genetic Linkage and Homogeneity

From: Julian Faraway
Department of Statistics
University of Michigan
Ann Arbor, Michigan 48109-1027, U.S.A.

To the Editor of Biometrics:

Lemdani and Pons (1995) discuss the problem of testing for genetic linkage when the data consist of linked and unlinked families in proportions λ and $1-\lambda$. The recombination fraction θ is $1/2$ when there is no linkage and less than $1/2$ when there is. They construct a likelihood ratio test statistic to test the null hypothesis that $\theta = 0$ against the alternative that $\theta < 1/2$ with $\lambda \in [\gamma, 1]$, and show that it has the asymptotic null distribution $1/2\delta_{\{0\}} + 1/2\chi_1^2$, which differs considerably from the result I obtained in Faraway (1993). The primary reason for this difference lies not in any mathematical error by either party but in Lemdani and Pons' specification that $\lambda \in [\gamma, 1]$ for some $\gamma > 0$ compared to mine that $\lambda \in (0, 1]$. When λ is allowed to lie in $(0, 1]$ and when the sample size becomes large, its maximum likelihood estimate $\hat{\lambda}$ is 1 about half the time and very small about half the time. If the sample size is sufficiently large, $\hat{\lambda}$ will be less than any fixed γ about half the time. Putting aside the mass $1/2$ at 0 due to the one-sided test, this explains why I find the limit distribution to be approximately the maximum of two independent χ_1^2 variates, while Lemdani and Pons find it to be χ_1^2, because if λ is restricted to $[\gamma, 1]$, $\hat{\lambda}$ will always be 1 as the sample size becomes large.

With small samples, it will not be possible to specify γ appropriately. If it is made very small, then their test statistic will be unaffected by the restriction that $\lambda \in [\gamma, 1]$. My own empirical investigations and those of previous authors reveal that $1/2\delta_{\{0\}} + 1/2\chi_1^2$ would then form a poor approximation to the null distribution.

A further point of difference is that I maximize the likelihood over $\theta \in [0, 1]$ and then simply accept the null if $\hat{\theta} \geq 1/2$. For some small samples it can happen that $\hat{\theta} \geq 1/2$ but the one-sided test statistic (where we restrict $\theta \in [0, 1/2]$ and $\lambda \in (0, 1]$) is strictly positive, even statistically significant. However, such cases are unusual, and asymptotically the one-sided test statistic and my version are the same. Contrary to Lemdani and Pons' claim, the root cause of the difficulty in finding the null distribution for these test statistics is that λ is unidentifiable under the null. Of course, my result is only approximate and the argument is heuristic. Chernoff and Lander (1995) have the definitive results.

REFERENCES

Chernoff, H. and Lander, E. (1995). Asymptotic distribution of the likelihood ratio test that the mixture of two binomials is a single binomial. *Journal of Statistical Planning and Inference* **43**, 19–40.

Faraway, J. (1993). Distribution of the admixture test under heterogeneity. *Genetic Epidemiology* **10**, 75–83.

Lemdani, M. and Pons, O. (1995). Tests for genetic linkage and homogeneity. *Biometrics* **51,** 1033–1041.

The authors replied as follows:

In Lemdani and Pons (1995) we compared two statistics which can be viewed as likelihood ratio test statistics for no linkage against two distinct restricted alternatives. Of course, we agree with J. Faraway concerning the definitive result about the likelihood ratio test that Chernoff and Lander (1995) derived from a judicious reparameterization. The null asymptotic distribution of the LR statistic is therefore $1/2\delta_0 + 1/2 \sup_\phi T^2(\phi)$, where T is a centered Gaussian process with a known covariance and $\phi = 2\theta - 1$ is the new parameter. In Lemdani and Pons (1996b) we rigorously proved this result and, using other reparameterizations, we extended it to the unknown phase case and to a test for linkage between a marker and a single gene against the alternative of two genes linked with the same marker (Chernoff and Lander, 1995, problem 4). We also used the same approach to study the asymptotic null distribution of the likelihood ratio statistic for a mixture of p distributions of the same form against an alternative of $p + q$ distributions (Lemdani and Pons, 1996a).

Generally, the recombination fraction θ is assumed to be in $[0, 1/2]$ since we are only interested in finding markers close to some gene of interest. However, if the true value of θ is close to $1/2$ and for small sample sizes, maximizing the likelihood without the restriction $\theta \in [0, 1/2]$ can clearly lead to an estimated value $\hat{\theta}$ greater than $1/2$. If the parameter space is extended to $[0, 1]$ and if we want to test $\theta = 1/2$, under the null hypothesis the LR statistic converges to the maximum of two squared truncated Gaussian processes instead of a single squared truncated Gaussian process as above (Pons and Lemdani, 1995, unpublished manuscript).

REFERENCES

Lemdani, M. and Pons, O. (1995). Tests for genetic linkage and homogeneity. *Biometrics* **51,** 1033–1041.
Lemdani, M. and Pons, O. (1996a). Likelihood ratio tests for mixture models. *C. R. Acad. Sci. Paris I* **322,** 399–404.
Lemdani, M. and Pons, O. (1996b). Likelihood ratio tests for genetic linkage. *Statist. Probab. Let.,* in press.

Mohamed Lemdani and Odile Pons
Biométrie
Institut National de la Recherche Agronomique
78352 Jouy-en-Josas, France