$$P(G) = \left\{X \mid X = (x_1, \cdots, x_n),\ x_i \geqq 0,\ \sum_{i=1}^{n} \alpha_{ij} x_i \leqq 1,\ 1 \leqq j \leqq m\right\}.$$

The polytopes $P(G)$ were introduced in [3] and their basic properties investigated. It was shown that the number of facets of $P(G)$ is $n+c(G)$ ($c(G)$- number of cliques in G) and that distinct graphs of $\Gamma(2)$ yield distinct polytopes. Since the number of n-polytopes with $n+2$ facets is $c(n, 2)$ (Grünbaum [1], p. 424), every n-polytope with $n+2$ facets is a $P(G)$ for an appropriate G. In the other extreme, Moon and Moser [2] have shown that the largest possible number of cliques in a graph on n vertices is $f(n)$, where:

$$f(n) = \begin{cases} 3^{n/3} & n = 0 \pmod 3 \\ 4 \cdot 3^{[n/3]-1} & n = 1 \pmod 3 \\ 2 \cdot 3^{[n/3]} & n = 2 \pmod 3). \end{cases}$$

A full characterization of graphs with $f(n)$ cliques is also given in [2], yielding

$$c(n, f(n)) = 1.$$

References

1. B. Grünbaum, Convex Polytopes, Interscience, London, 1967.
2. J. W. Moon, L. Moser, On cliques in graphs, Israel J. Math, 3 (1965) 11, 23–28.
3. M. Rosenfeld, On certain polytopes associated with graphs, Proc. of the Conference on Combinatorics and Graph Theory, L.S.U., 1970 (to appear).

CLASSROOM NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to David Drasin, Division of Mathematical Sciences, Purdue University, Lafayette, IN 47907.

MAXIMUM LIKELIHOOD AND SUFFICIENT STATISTICS

D. S. MOORE, Purdue University

A theorem common to many statistics texts asserts that a maximum likelihood estimator must be a function of any sufficient statistic. While true in principle and true in most standard problems, this assertion is false as stated. In this note I want to clarify this situation with some comments and an example.

Let $X_1, \cdots, X_n$ be a random sample from a distribution having pdf $f(x \mid \theta)$, where θ is a real parameter lying in a parameter space Ω but otherwise unknown. Then a *maximum likelihood estimator* (MLE) of θ is any function $\hat{\theta} = \hat{\theta}(x_1, \cdots, x_n)$ which maximizes the joint pdf $\prod_{i=1}^{n} f(x_i \mid \theta)$ for each set of x_i in the sample space. A *sufficient statistic* is intuitively a function $t(X_1, \cdots, X_n)$ of the sample which compresses the data without losing information about θ.

This idea is formalized by requiring that the conditional distribution of the sample given the value of t not depend on θ. It seems reasonable that statistical procedures should be based on the values of a sufficient statistic.

The "proof" that any MLE must be a function of any sufficient statistic is extremely simple: if t is sufficient, the factorization criterion for sufficiency states that a factorization of the form

$$\prod_{i=1}^{n} f(x_i \mid \theta) = g(\theta, t)h(x_1, \cdots, x_n)$$

is possible. The $\hat{\theta}$ maximizing the joint pdf for fixed $x_1, \cdots, x_n$ therefore depends on the x_i only through the value of t. This theorem and proof are given in some of the best undergraduate texts (e.g., Hogg and Craig [3], p. 256). The simplicity of the proof leads authors of more advanced texts to relegate the result to a remark (Rao [4], p. 291) or a problem (Ferguson [2], p. 125).

A correct statement of the theorem is: *If t is sufficient for θ and a unique MLE $\hat{\theta}$ of θ exists, $\hat{\theta}$ must be a function of t. If any MLE exists, an MLE $\hat{\theta}$ can be chosen to be a function of t.*

This reformulation of the theorem should point out the defect in the "proof." It follows from the factorization criterion that the *set* of values of θ maximizing the joint pdf depends on the sample only through the value of t. If the maximizing value of θ is not unique, we must choose a particular MLE $\hat{\theta}$ by selecting a value from the maximizing set. This can, but need not, be done in a manner independent of the sample.

The following example illustrates these points. Suppose $X_1, \cdots, X_n$ have the uniform distribution on $(\theta-\frac{1}{2}, \theta+\frac{1}{2})$. Then it is easy to see that $t=(\min X_i, \max X_i)$ is a sufficient statistic and that any θ satisfying

$$\max x_i - \tfrac{1}{2} \leqq \theta \leqq \min x_i + \tfrac{1}{2}$$

maximizes the joint pdf for given $x_1, \cdots, x_n$ (Hogg and Craig [3], p. 255). We can therefore choose an MLE which is a function of t, for example

$$\hat{\theta} = \tfrac{1}{2}(\min X_i + \max X_i).$$

But we can also choose an MLE which is not a function of t, such as

$$\hat{\theta} = (\max X_i - \tfrac{1}{2}) + (\cos^2 X_1)(\min X_i - \max X_i + 1).$$

This example is slightly pathological in that the dimension of the sufficient statistic exceeds that of the parameter. No real sufficient statistic exists in the example, but in most standard statistical problems a real parameter has a real sufficient statistic. In such cases it can be shown that (under mild regularity conditions) the MLE is unique and is therefore a function of the sufficient statistic. The proof depends on theorems asserting that existence of a continuous real function t which is sufficient implies that $f(x \mid \theta)$ is a single parameter exponential family (see Brown [1]). The simple "proof" discussed above remains inadequate.

References

1. L. Brown, Sufficient statistics in the case of independent random variables, Ann. Math. Statist., 35 (1964) 1456.

2. T. S. Ferguson, Mathematical Statistics: A Decision Theoretic Approach, Academic Press, New York, 1967.

3. R. V. Hogg, and Allen Craig, Introduction to Mathematical Statistics, 3rd ed., Macmillan, New York, 1970.

4. C. R. Rao, Linear Statistical Inference and its Applications, Wiley, New York, 1965.

A REGULAR SPACE ON WHICH EVERY CONTINUOUS REAL-VALUED FUNCTION IS CONSTANT

T. E. GANTNER, University of Dayton

The well-known examples, due to Hewitt [2] and Novák [3], of the type of spaces mentioned in the title of this article are not easy to present in an introductory level course of topology because their constructions depend heavily on the theory of cardinal numbers. More recently, Herrlich [1] outlined an easier construction of such a space, but his construction also involves many properties of cardinal numbers. The purpose of this note is to outline the construction of such a space that avoids all cardinality arguments except the distinction between countable and uncountable sets.

The space with which we begin the construction is a regular space Q that has two points p^- and p^+ such that $f(p^-)=f(p^+)$ for every continuous real-valued function f on Q. Thomas [4] recently gave a simple geometric construction of such a space using only the distinction between countable and uncountable sets. The construction of such a space is the most difficult part of Herrlich's paper, and it is the part that depends on cardinal numbers. One can now proceed *verbatim* with the construction as outlined in Herrlich's paper. For completeness we include these steps:

PROPOSITION: *For any regular space Z, there exists a regular space $Q(Z)$ such that Z is imbedded as a subspace of $Q(Z)$ and every continuous real-valued function of $Q(Z)$ is constant on Z.*

To obtain $Q(Z)$, one first constructs a topology on the product set $Z\times Q$ by declaring a subset $V\subset Z\times Q$ to be open if the following are true:

1. If $(z, x)\in V$, then there is a neighborhood U of x in Q such that $\{z\}\times U\subset V$.

2. If $(z, p^+)\in V$, then there is a neighborhood U of z in Z such that $U\times\{p^+\}\subset V$.

With this topology $Z\times Q$ is regular and Z is homeomorphic to the line $Z\times\{p^+\}$ in $Z\times Q$. It should be pointed out that the neighborhoods on the upper edge $Z\times\{p^+\}$ of the square $Z\times Q$ are somewhat complicated; one may picture them as appearing like "ragged-edged combs." The space $Q(Z)$ is now obtained by identifying all the points of the line $Z\times\{p^-\}$ in $Z\times Q$.