

NONPARAMETRIC STATISTICAL INFERENCE

by

Jean D. Gibbons

New York : McGraw-Hill Book Company, 1971, 306 sayfa, ciltli

«Parametreye ilişkin olmayan (nonparametric) istatistik» veya diğer ismi ile «olasılık dağılımı ile bağımlı bulunmayan (distribution-free) istatistik» konusunda istatistikçiler, uzun yıllar, doyurucu ve birleştirici kuramsal (teorik) bir metnin özlemini çekmişlerdir¹. Bu sıralar, bahse konu istatistiği oluşturan konular, ilk yayınlandıkları bilimsel ve mesleki dergilerde parça parça izlenebiliyordu. Ara sıra basılan metinler ise, ya yüzeysel veya uygulamaya yönelik (Siegel, 1956)² kalmakta, ya da çok özel konuları içermekteydi (Gumbel, 1958) (Kendall, 1962). Son yedi, sekiz yıldır ortaya çıkan kitaplara gelince; gerek içerik, gerek sayıları bakımından, bu son yayınlar dikkati çekmektedir: (Noether, 1967), (Hájek ve Šidák, 1967), (Bradley, 1968), (Hájek, 1969) (Gibbons, 1971), (Puri ve Sen, 1971), (Hollander ve Wolfe, 1973).

Bu son yapıtlardan bazılarının matematiksel yaklaşımı bir kesim okuyucu ve öğrencinin gözünü korkutacak derecede ileri düzeydedir: (Hájek ve Šidák, 1967), (Puri ve Sen, 1971). Bazılarının içeriği de uygulamaya göre düzenlenmiş ve böylece de teorik istatistikçilerin beğenisini kazanamayacak bir biçime bürünmüştür: (Hollander ve Wolfe, 1973). Bu yüzden, özellikle tek değişkenli, parametreye ilişkin olmayan istatistik kuramını işleyen iki kitap göze batmaktadır: (Hájek, 1969), (Gibbons, 1971).

Bunlardan günümüze daha yakın bir tarihte kaleme alınmış olan sonuncusu üzerinde durmak yararlı olacaktır:

Kitabın önsözünde Bayan Gibbons tarafından belirtildiği gibi, parametreye ilişkin olmayan istatistik konusundaki kitaplar, genellikle

- (1) Bu iki isim gerçekte özdeş konuları belirtmemektedir. "Olasılık dağılımı ile bağımlı bulunmayan istatistik" ismi altında toplanan konulara da, nedense, "parametreye ilişkin olmayan istatistik" ismi verilegelinmektedir.
- (2) Bu kitap (Savage, 1957) de eleştirilmiştir.

ie, ya el kitabı yüzeyselliğini aşamamakta, ya aşırı matematiksel bir havaya bürünmektedir. Yazarın işaret ettiğine göre, kitap, bu iki aşırı yaklaşımı uzlaştırmak amacı ile kaleme alınmıştır. Bunda da bir dereceye kadar başarıya ulaşılmıştır.

Gerçekten, parametreye ilişkin olmayan istatistik, uzun süre, bir çeşit teknik hüneri ve beceri gösterisi, bir yol ve yordam şekli olarak görülmüş; bu hüneri ve becerinin arkasında herkesçe anlaşılabilir köklü bir kuram bulunabileceği dikkate alınmamıştır. Bu durumun ortaya çıkardığı bir sakınca da, kuşkusuz, iyi anlaşılmamış bir kuramın yanlış uygulamalara yol açması ve yanılgılara neden olmasıdır. Öte yanda, bölük pörçük ilgili bilimsel ve mesleki dergilere saçılmış olan kuramsal tartışmalar ise, meslekten gelmeyen araştırmacıları yığınla sürükleyecek derecede meslekçi (profesyonel) ve dağınıktır.³ Gibbons'ın yaptığı şey, işte bu dağınık okuma geçrelerini, tek tip bir yorumla bir araya toplaması ve izlenmesi kolay bir kuramsal yaklaşımla açıklamasıdır. Bu açıklamalar, (Parzen 1960) veya (Mood ve Graybill, 1963)⁴ gibi standard bir ders kitabı düzeyindeki olasılık ve istatistiksel çıkarıma (statistical inference) bilgisiyle rahatlıkla izlenebilmektedir.

Konular kitapta kuramsal niteliklerine göre düzenlenmiştir. İlerideki eleştirmelerimizi saklı tutmak koşulu ile, bu düzenlemenin, bir bakıma, geleneksel konu ayırımına uygun düştüğünü belirtmekte yarar vardır. Böylece, yirmi bir sayfalık bir giriş tartışmasından sonra, ikinci bölümde **sıra istatistikleri** (order statistics) tartışılmıştır. Üçüncü bölüm, **izleyiş (veya zincirleme) kuramı** diye çevirebileceğimiz "theory of runs" konusuna ilişkindir. Dördüncü ve beşinci bölümler, sırası ile, **uyumun doğruluğu** (goodness-of-fit) ve **derece (aşama) istatistikleri** (rank-order statistics) konularını ele almaktadır.

Bundan sonraki iki bölüm, değişik bir konu görünümünde olan, tek örneklem ve çift örneklem testlerini ele almaktadır. Ona göre, **cebirsel işaret testleri** (sign test ve signed-rank test) beşinci; Wald-Wolfowitz izleyiş testleri, medyan (ortanca) testi ve Mann-Whitney U testi gibi çeşitli **çift örneklem testleri** de altıncı bölümdedir. Bilindiği gibi, bu testler parametreye ilişkin (parametric) istatistikte sık sık rastladığımız normal dağılım, t-dağılımı ve F-dağılımı testlerine karşın bulunmaktadır.

Sekizinci, dokuzuncu ve onuncu bölümler ise, **doğrusal derece**

(3) Örneğin, (Savage, 1962) deki bibliyografyaya bakılabilir.

(4) Bu yapıt türkçeleştirilmiş ve İstanbul Teknik Üniversitesi tarafından yayınlanmıştır.

istatistiklerine (linear rank statistics) ilişkindir. Bu başlık altında tartışılan konular Wilcoxon testi ile çeşitli konum (location) ve ölçek (scale) sorunlarıdır. Onbirinci bölüm, altı ve yedinci bölümlerde tartışılan konunun ikiden fazla örnekleme kapsayacak şekilde genişletilmesidir. Onikinci ve onüçüncü bölümlerde de, ikili ve çoklu örneklem için parametreye ilişkin olmayan çeşitli **ortasım ölçüleri** (measures of association) ele alınmıştır. Bilindiği gibi korelasyon katsayısı buraya girmektedir. Parametreye ilişkin olmayan istatistikte önemli bir yeri olan **asemptotik görelî etkinlik** (asymptotic relative efficiency) son bölümde tartışılmaktadır.

Konuların bu şekilde düzenlenmesinin bazı sakıncaları da bulunmaktadır. Örneğin, medyan testi sıra istatistikleri konusunu ilgilendirir ve bu nedenle ikinci bölüme girmesi gerekirken, kitabın yedinci bölümünde yer almıştır. Bunun gibi, gene medyanla yakın ilişkisi olan cebirsel işaret testinin altıncı bölümde tartışıldığı görülmektedir. Böylece, kitabın ikinci bölümünden altıncı ve yedinci bölümlerine iletme (reference) yapılmaktadır.

Bu kitapta tartışılan her konunun başlı başına ayrı bir kitaba eşdeğer olduğu anımsanırsa, Gibbons'ın eserinde her konunun ayrıntıları ile incelenemeyişinin nedenini anlamak güç olmamaktadır. Örneğin, ikinci bölümün içeriğini (David, 1970), beşinci bölüme ilişkin konuları (Kendall, 1962) ve son bölümlere ilişkin konuları da (Puri ve Sen 1971) gibi ayrı kitaplarda izlemek olanağı bulunmaktadır. Bu yüzden, örneğin, aşırı uç değerleri (outliers) ve kayma (slippage) gibi konular ikinci bölümde; izleyişin uzunluğu (length of runs) konusu üçüncü bölümde ve çoklu analizle (multivariate analysis) rastlantısal yöneşim (stochastic convergence) son bölümlerde ayrıntıları ile işlenememiştir.

Bununla beraber, ele alınan konular, bir matematikçiden beklenen kesinlik ve yalınlıkla biçimlendirilmiş, aşırı matematik isteyen birkaç teorem dışındaki bütün teoremlerin kanıtı gösterilmiştir. Bu nedenle, dikkatli bir okuyucu, çeşitli sorunlar karşısında, türlü kuramsal yaklaşımla neler yapabileceğine değgin açık fikir edinmekte; bu arada, başvurulabilecek kestirme yöntem ve hazırlanmış tabloları nerelerde bulabileceğini öğrenmek olanağına kavuşmaktadır.

Yukarıda sayılan nedenlerle, Bayan Gibbons'ın kitabını, hiç olmazsa, orta düzeydeki kuramsal istatistik öğretimine olumlu bir katkı şeklinde değerlendirmek olasıdır.

Yalçın TUNCER

Orta Doğu Teknik Üniversitesi
Uygulamalı İstatistik Bölümü