

İSTATİSTİKSEL GÜVEN VE MANİDARLIK KAVRAMLARI (*)

Çeviren : Doç. Dr. İlhan AKHUN

Robert E. CHANDLER

Son yıllarda en azından üç ayrı kitap eleştirmeni istatistiksel güven ve manidarlık kavramları ile ilgili olarak psikolojik literatürde halihazırdaki karışıklık konusunu tartışmıştır (2, 7, 9). Bu karışıklık semantik bir sorun olarak kısmen açıklanabilmekle beraber, bu psikologun bu iki kavramı daha yakından incelemesini ve psikolojik literatürü büyük ölçüde kullanan diğer disiplinlerin başlangıç öğrencileri ve bireylerinin yararlanması için eski terminolojiyi benimsemesini gerektirir.

Güven ve Güven Katsayıları

Genellikle aralık hesaplanması tartışmalarında kullanılan bir kavram olan güven, bir kimsenin bir örneklem işlemi ile saptanan bir aralığın belli bir parametreyi kapsamasına ilişkin bir ifadeye duyduğu itimattır. Bir kimse genellikle bu itimadı istatistiksel olarak her aralığa $1 - \epsilon$ olarak yazılabilen bir güven katsayısı veya güven olasılığı vermekle belirtebilir. Burada, $0 \leq p \leq 100$ için $\epsilon = p/100$ dür ve p genellikle küçük bir sayı olarak alınabilir (1, 3, 6, 8). Örneğin, $p=5$ ise güven katsayısı .95 olacak ve bir kimse bu katsayıya eşlik eden aralığa % 95 güven aralığı olarak tanımlayacaktır.

Güven katsayısı çoğu kez aşağıdaki gibi yorumlanır: Eğer bir kimse N elemanı olan bir evrenden K büyüklüğündeki örneklem almış (doğal olarak $K < N$ olur) ve her bir örneklemden evrenin belli bir parametresi için % 95 güven aralığını saptamış olsaydı, bu durumda uzun sürede bu aralıkların toplamının % 95'i ilgili parametreyi gerçekten kapsardı ve bun-

(*) Robert E. Chandler, «The Statistical Concepts of Confidence and Significance,» *Psychological Bulletin*, 54, 429 - 430.

ların yaklaşık olarak % 100 e'ü veya % 5'i kapsamazdı (3). Bu yorum doğrudur, fakat tabiatıyla $\binom{N}{K}$ nın daha büyük bir sayı olmasını kabul eder.¹

Manidarlık ve Manidarlık Düzeyleri

Güven kavramının aksine, manidarlık kavramı hipotezleri test için kullanılır. Burada bir kimse bir ifade kullanır; örneğin, buradaki tartışmada bundan sonra H ile gösterilecek olan bir hipotezi ifade eder ki, bu doğru veya yanlış olabilir ve daha sonra bunu kabul veya red ederek bu H konusunda eyleme geçer. Açıkcası, bir H testinin sonucu olarak aşağıdaki durumlardan herhangi biri olası bir sonuçtur: (a) yanlış bir H'in reddedilmesi; (b) doğru bir H'in kabul edilmesi; (c) doğru bir H'in reddedilmesi; veya (d) yanlış bir H'in kabul edilmesi. (a) ve (b) ile belirtilen eylemlerin arzu edilen birer eylem olduğu oldukça açıktır, oysa (c) ve (d) bir hatayı ortaya çıkarır — c Tür I hata veya birinci tür hata olarak bilinir, d'ye ise Tür II hata veya ikinci tür hata adı verilir (4, 8).

Bir kimse, bir H'i test ederken, c eylemine geçmesinin olasılığı manidarlık düzeyi olarak tanımlanır, bunu da α ile göstereceğiz (8). α nın genellikle ϵ gibi aynı önem derecesine sahip olmasına karşın, α ve ϵ verdikleri bilginin miktarı bakımından farklılaşır. ϵ tamamen aralık hesaplanmasında «yanılmış olma» hakkında bilgi verirken, α yalnızca c ile tanımlanan bir eylem gibi çok özel olan bir hata türü hakkındaki bilgiyi verir. Burada α ve ϵ arasındaki farkı vurgulamak için, birisinin yalnızca bir H testinde yapılan diğer hata türünü incelemesi gerekir.

Bu amaç için, β eylem d'nin olasılığını, örneğin, Tür II hatayı temsil etsin; bu durumda, tanıma göre $1 - \beta$ istatistiksel testin gücü veya eylem a'nın meydana gelmesinin olasılığı olarak bilinir (4, 8). Psikolojik istatistik ile ilgili ders kitaplarının bir testin gücü konusuna büyük bir önem vermemesine karşın, güç bir kimsenin bir H üzerinde karar vermek için temel olarak kullandığı istatistiksel test için önemli bir kavramdır. Eğer böyle değilse, bir H'i % 5 manidarlık düzeyinde test etmek için, bir

1. Buradaki $\binom{N}{K}$ işareti N obje arasından K objeyi seçmenin değişik biçimlerini göstermek için kullanılan bir simgedir.

kimse içinde 100 boncuk - 95 beyaz ve 5 kırmızı - bulunan bir kutudan yansız (random) olarak bir boncuk çeker ve kırmızı bir boncuk çektiğinde H'in reddedileceği konusundaki kararı benimser. Böyle bir test ile, bir kimse yalnız α 'nın değil, fakat $(1 - \beta)$ 'nin de her zaman .05'e eşit veya $\beta = .95$ olduğunu açıkça görebilir. β 'nin bu büyük değeri bir kimsenin boncuk kutusu testini kullanmasına engel olur. β 'nin çok iyi bir tartışması ve bunun bir kimsenin test ettiği karşıt H'e olan ilişkisi için, okuyucunun Dixon ve Massey'in kitabına başvurması önerilir (4, s. 244 - 261).

Özet ve Tartışma

Güven ve manidarlık kavramlarının karmaşıklığı psikolojik literatürde o kadar belirgin bir duruma gelmiştir ki, bir kimsenin psikolojik araştırma raporlarında belli bir deneme sonuçlarının, örneğin % 5 «güven düzeyinde,» manidar olduğunu belirten ifadeleri okuması tipiktir.

Bu karmaşıklık bir kimsenin bir güven aralığını bir manidarlık testi (örneğin, bak. 5, s. 241) olarak kullanmasından ve böyle yapmakla, aceleci fakat yanlış olarak, iki kavram arasında bir fark olmadığı sonucuna varmasından ortaya çıkabilir.

Aralık hesaplanması ve hipotezlerin test edilmesi ile ilgili alanlarda yapılan istatistiksel hataların olasılıklarını açıklamak için açık bir terminolojiye gereksinim duyulduğundan, güven kavramı ile kullanılan testin özelliğine bakmaksızın bir H'in test edilmesi arasında bir ilişkinin hiç bir zaman kurulmaması gerekir.

KAYNAKLAR

1. Anderson, R. L., and Bancroft, T. A. *Statistical Theory in Research*. New York: McGraw-Hill, 1952.
2. Chandler, R. E. A review of Guilford's *Fundamental Statistics in Psychology and Education*. (3rd ed.) *Personnel Psychology*, 1957, 10, 272-273.
3. Cramer, H. *Mathematical Methods of Statistics*. Princeton: University Press, 1951.
4. Dixon, W. J., and Massey, F. J., Jr. *Introduction to Statistical Analysis*. (2nd ed.) New York: McGraw-Hill, 1957.

5. Edwards, A. L. **Statistical Methods for the Behavioral Sciences**. New York : Rinehart, 1954.
6. Hoel, P. G. **Introduction to Mathematical Statistics**. (2nd ed.) New York : Wiley, 1954.
7. Milton, T. E. A review of Edwards' **Statistical Methods for the Behavioral Sciences**. **Journal of American Statistical Association**, 1956, 51, 382.
8. Mood, A. M. **Introduction to the Theory of Statistics**. New York : McGraw-Hill, 1950.
9. Walker, H. M. A review of Adam's **Basic Statistical Concepts**. **Educational Psychological Measurement**, 1956, 16, 554 - 557.