آزمون گِری برای نرمال بودن

علی عمیدی*

نمونهٔ تصادفی X_1، X_2، …، X_n را از جامعهٔ متغیر تصادفی X به دست آورده‌ایم. می‌خواهیم آزمون کنیم که این نمونه از جامعه‌ای نرمال آمده است یا نه. بدیهی است آزمونهای کلاسیک خی دو و کولموگوروف را می‌توان بر حسب شرایط موجود و با توجه به مزایا و معایب هر کدام به کاربرد. در زیر، به صورتی خلاصه، آزمون دیگری را که بنیانگذار آن گری ($Geary$) است متذکر می‌شویم.

این آزمون بر پایهٔ آمارهٔ زیر است:

$$G = \frac{\sum_{i=1}^{n} |X_i - \bar{X}_n|}{\left[n \sum_{i=1}^{n} (X_i - \bar{X}_n)^2\right]^{\frac{1}{2}}} = \frac{\frac{1}{n}\sum_{i=1}^{n} |X_i - \bar{X}_n|}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}}$$

در واقع، آمارهٔ آزمون، نسبت میانگین قدر مطلق انحراف متغیرها از میانگین نمونه به انحراف معیار نمونه است. گری صدکهای توزیع G را تحت فرض نرمال بودن جامعه حساب کرده است. فرض نرمال بودن جامعه در سطح معنادار بودن α وقتی رد می‌شود که مقدار تجربی G بیرون بازهٔ $(G_{\frac{\alpha}{2}}, G_{1-\frac{\alpha}{2}})$ باشد. مقادیر دو سربازه از جدول صدکهای توزیع G به دست می‌آیند. در زیر خلاصه‌ای از این جدول را آورده‌ایم. یادآور می‌شویم که وقتی بافتنمای داده‌ها پخ باشد، حساسیت این آزمون بیشتر است.

مثال . نمونه‌ای به اندازهٔ ۲۰ از توزیع نرمال $N(2, 1)$ انتخاب کرده‌ایم. مرتب شدهٔ مشاهدات نمونه‌ای در زیر آورده شده است:

۰٫۷۹۰	۰٫۸۴۳	۱٫۲۸۶	۱٫۳۳۱	۱٫۳۸۳
۱٫۴۱۱	۱٫۷۵۷	۱٫۸۷۴	۱٫۹۸۵	۲٫۰۰۶
۲٫۰۰۹	۲٫۳۵۵	۲٫۴۸۱	۲٫۵۶۰	۲٫۹۸۰
۳٫۱۲۹	۳٫۲۸۴	۳٫۲۸۷	۳٫۳۲۸	۳٫۵۳۲

به راحتی می‌توان دید که

$$\sum_i |x_i - \bar{x}_m| = 14.620 \quad , \quad \bar{x}_n = 2.181$$

$$G = 0.866 \quad , \quad \sum_{i=1}^{20} (x_i - \bar{x}_n)^2 = 14.232$$

برای سطح معنادار بودن ۰٫۱، با توجه به جدول، بازهٔ پذیرش فرض نرمال بودن به صورت (۰٫۷۲۹, ۰٫۸۷۹) است. حال چون $(G=)$ ۰٫۸۶۶ درون بازهٔ بالا می‌افتد، در سطح معنادار بودن ۰٫۱، فرض نرمال بودن جامعه رد نمی‌شود.

n	$E(a)$	%۱	%۵	%۹۵	%۹۹
۱۰	۰٫۸۱۹	۰٫۶۶۴	۰٫۷۱۴	۰٫۹۱۱	۰٫۹۴۰
۱۵	۰٫۸۱۳	۰٫۶۸۰	۰٫۷۲۲	۰٫۸۹۲	۰٫۹۱۸
۲۰	۰٫۸۰۹	۰٫۶۹۳	۰٫۷۲۹	۰٫۸۷۹	۰٫۹۰۳
۲۵	۰٫۸۰۶	۰٫۷۰۲	۰٫۷۳۵	۰٫۸۷۰	۰٫۸۹۲
۳۰	۰٫۸۰۵	۰٫۷۱۰	۰٫۷۳۹	۰٫۸۶۴	۰٫۸۸۴
۳۵	۰٫۸۰۴	۰٫۷۱۶	۰٫۷۴۳	۰٫۸۵۹	۰٫۸۷۸
۴۰	۰٫۸۰۳	۰٫۷۲۰	۰٫۷۴۶	۰٫۸۵۵	۰٫۸۷۳
۴۵	۰٫۸۰۲	۰٫۷۲۵	۰٫۷۴۹	۰٫۸۵۱	۰٫۸۶۹
۵۰	۰٫۹۰۲	۰٫۷۲۸	۰٫۷۵۱	۰٫۸۴۹	۰٫۸۶۵
۶۰	۰٫۸۰۱	۰٫۷۳۴	۰٫۷۵۵	۰٫۸۴۴	۰٫۸۶۰
۷۰	۰٫۸۰۱	۰٫۷۳۸	۰٫۷۵۸	۰٫۸۴۱	۰٫۸۵۶
۸۰	۰٫۸۰۰	۰٫۷۴۲	۰٫۷۶۰	۰٫۸۳۸	۰٫۸۵۲
۹۰	۰٫۸۰۰	۰٫۷۴۵	۰٫۷۶۲	۰٫۸۳۶	۰٫۸۴۹
۱۰۰	۰٫۸۰۰	۰٫۷۴۸	۰٫۷۶۴	۰٫۸۳۵	۰٫۸۴۶
۲۰۰	۰٫۷۹۹	۰٫۷۶۲	۰٫۷۷۴	۰٫۸۲۵	۰٫۸۳۲
۳۰۰	۰٫۷۹۹	۰٫۷۶۹	۰٫۷۷۸	۰٫۸۱۹	۰٫۸۲۶
۴۰۰	۰٫۷۹۸	۰٫۷۷۲	۰٫۷۸۱	۰٫۸۱۶	۰٫۸۲۲
۵۰۰	۰٫۷۹۸	۰٫۷۷۵	۰٫۷۸۲	۰٫۸۱۴	۰٫۸۲۰
۶۰۰	۰٫۷۹۸	۰٫۷۷۷	۰٫۷۸۴	۰٫۸۱۳	۰٫۸۱۸
۸۰۰	۰٫۷۹۸	۰٫۷۸۰	۰٫۷۸۶	۰٫۸۱۰	۰٫۸۱۵
۱۰۰۰	۰٫۷۹۸	۰٫۷۸۲	۰٫۷۸۷	۰٫۸۰۹	۰٫۸۱۳
∞	۰٫۷۹۷۹				

*ـ دکتر علی عمیدی، گروه آمار، دانشگاه شهید بهشتی