

آزمون های معنی داری

با کمک آزمون های معنی داری می توان پی برد که کدام تفسیر درست است. منطق این آزمون ها ساده است. اگر دو متغیر در جمعیت فاقد رابطه باشند احتمال این که نمونه تصادفی ما بیان گر رابطه ای بین این دو متغیر باشد چقدر است؟ (یعنی احتمال دقیق نبودن نمونه تصادفی ما چقدر است؟). به عنوان مثال اگر صدبار نمونه گیری تصادفی انجام دهیم، احتمال این که یکی از آن ها نمونه غیر معرفی باشد، یعنی بیان گر رابطه ای باشد که واقعا در جمعیت وجود ندارد چقدر است؟ معمولا گفته می شود آنجا که احتمالا از هر صد نمونه بیش از پنج نمونه بیان گر رابطه ای باشند که ناشی از خطای نمونه گیری است، احتمال نادرست بودن نمونه بالاست. چه بسا نمونه خاص ما یکی از پنج نمونه باشد! در نتیجه باید گفت به احتمال زیاد رابطه مشاهده شده ناشی از خطای نمونه گیری است و فرض فقدان رابطه در جمعیت واقعی صحیح است.

گروهی از پژوهش گران محتاط ترند و معتقدند آنجا که بیش از یکی از صد نمونه بتواند برحسب تصادف رابطه ای به شدت رابطه مشاهده شده ایجاد کند، آن گاه احتمال خطا زیاد است. اما اگر دریابیم که صرفا شمار بسیار اندکی از نمونه ها ممکن است رابطه مشاهده شده را ایجاد کنند می توانیم قبول کنیم که رابطه مشاهده شده در نمونه ما واقعی و منعکس کننده رابطه در جمعیت است.

از آن جا که هرگز صدبار نمونه گیری نمی کنیم، باید برآورد کنیم که اگر صد بار نمونه گیری می کردیم چقدر احتمال داشت که نمونه ما جزء یکی از نمونه هایی باشد که صرفا بر اساس شانس و تصادف بیان گر رابطه ای به شدت رابطه مشاهده شده در نمونه ما هستند.

با کمک نظریه احتمالات می توان احتمال واقعی نبودن رابطه مشاهده شده در نمونه را (یعنی احتمال این که ناشی از خطای نمونه گیری باشد) برآورد کرد. (در این جا به این نظریه نمی پردازیم و فقط متذکر می شویم که فرض این نظریه این است که از نمونه های تصادفی استفاده می کنیم). آزمون معنی داری آماری در واقع برآورد همین احتمال است. دامنه مقادیر این آزمون ها از (۰) تا (۱) است و آن ها را سطوح معنی داری می خوانند. معنای این ارقام چیست؟ گیریم سطح معنی داری ۰/۵۰ است. این بدان معناست که در ۵۰ نمونه از ۱۰۰ نمونه فقط بر اثر خطای نمونه گیری (شانس) رابطه ای به قوت رابطه ای که ما در نمونه مشاهده کردیم دیده می شود. در این صورت احتیاط حکم می کند رابطه مشاهده شده در نمونه خود را به احتمال زیاد واقعی ندانیم و در نتیجه فرض عدم رابطه در جمعیت رد نمی شود (تأیید می شود).

اگر سطح معنی داری ۰/۰۵ باشد بدان معناست که فقط پنج نمونه از هر ۱۰۰ نمونه برحسب تصادف به رابطه مشاهده شده در نمونه ما منجر می شود. اگر سطح معنی داری ۰/۰۱ باشد به معنی واقعی نبودن رابطه در یک نمونه از هر ۱۰۰ نمونه و اگر ۰/۰۰۱ باشد به معنای یک نمونه در ۱۰۰۰ نمونه است. پیداست هر چه سطح معنی داری پایین تر باشد، می توان اطمینان بیشتری به «واقعی» بودن رابطه مشاهده شده در نمونه داشت.

در اینجا به نحوه محاسبه این آزمون‌های آماری معنی‌داری نمی‌پردازیم. فرمول این‌ها در کتاب‌های آماری وجود دارد و با برنامه‌های کامپیوتری هم به راحتی می‌توان آن‌ها را حساب کرد. اما مسأله‌ای وجود دارد که برنامه‌های کامپیوتری پاسخگوی آن نیستند. اکثر برنامه‌های کامپیوتری سطح معنی‌داری را بین ۰/۰ تا ۱/۰۰ محاسبه می‌کنند؛ اما در چه سطحی فرضیه عدم رابطه در جمعیت (فرضیه صفر) رد می‌شود؟ معمولاً سطح معنی‌داری را ۰/۰۵ تا ۰/۰۱ را به عنوان مبنا در نظر می‌گیرند. اما این سطوح قراردادی و اختیاری هستند.

مسأله‌ای که در کاربرد سطح معنی‌داری ۰/۰۵ وجود دارد، سهولت رد فرض صفر (عدم رابطه) است: چه بسا فرض عدم رابطه در جمعیت رد شود (فرض وجود رابطه تأیید شود) در حالی که واقعاً رابطه‌ای وجود نداشته باشد. چنین اشتباهی را **خطای نوع اول** می‌خوانند و بیشتر در نمونه‌های بزرگ پیش می‌آید. از این رو بهتر است در نمونه‌های بزرگ سطح معنی‌داری ۰/۰۱ را به عنوان مبنا در نظر بگیریم. اما اگر همواره از سطح ۰/۰۱ استفاده کنیم ممکن است کار به **خطای نوع دوم** بکشد - یعنی سخت‌گیری بیش از حد و تصدیق فرضیه صفر در جایی که باید آن را رد کرد. احتمال چنین خطایی در نمونه‌های کوچک بیشتر است. طبق قاعده تجربی برای نمونه‌های کوچک از سطح معنی‌داری ۰/۰۵ و برای نمونه‌های بزرگ از سطح ۰/۰۱ یا کمتر استفاده می‌کنیم.

☑ نکته: سطح معنی‌داری در برنامه SPSS به صورت Sig گزارش می‌شود که در هنگام گزارش نتایج در پایان‌نامه‌ها و مقالات باید به صورت مقدار P (P-Value) گزارش شود.