

نقش کیفیت و ثبات داده‌های خام در اعتبار یافته‌های پژوهش

دکتر اسفندیار ستوده مرام* - دکتر شهناز روانشاد**

* منخصص اپیدمیولوژی دانشگاه علوم پزشکی و خدمات بهداشتی درمانی استان فارس

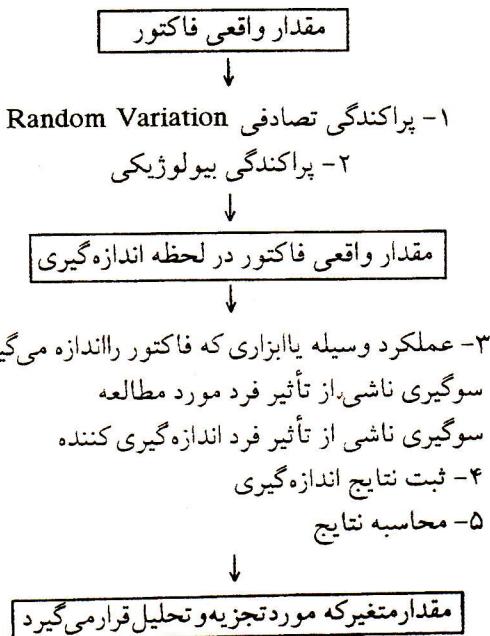
** منخصص تغذیه دانشگاه علوم پزشکی و خدمات بهداشتی درمانی استان فارس

کلید واژه‌ها: آمار / آمار و داده‌های عددی / پژوهش / مدل‌های آماری

مقدمه

کیفیت داده‌های خام آن، قبل از تجزیه و تحلیل‌های آماری، بالا یا بهتر از دیگری باشد.

کیفیت داده‌های خام از دو جنبه حائز اهمیت است، یکی مقدار اندازه‌گیری شده متغیر و برابر بودن با، یا نزدیک بودن آن به مقدار واقعی *True Value*، و دیگری ثبات *Consistency* داده‌ها. عوامل مختلفی می‌توانند موجب تفاوت بین مقدار اندازه‌گیری شده و مقدار واقعی یک متغیر بشوند. این عوامل در شکل شماره ۱ نشان داده شده است.



یکی از عوامل مهم و تعیین کننده اعتبار یافته‌های هر پژوهشی کیفیت داده‌های خام آن است. صرفنظر از عوامل دیگری مثل طراحی پژوهش و روش کار، نحوه انتخاب افراد مورد مطالعه و... که هریک بنویه خود از اهمیت زیادی برخوردار هستند، روشهای جمع آوری داده‌ها باید بگونه‌ای باشد که داده‌ها از کیفیت بالائی برخوردار باشند.

با اینکه هر محققی منتهای سعی خود را برای بالا بردن کیفیت داده‌های خام پژوهش بعمل می‌آورد، ولی در عین حال باید بطرق مناسب به خواننده گزارش خود نشان دهد که داده‌های جمع آوری شده، قبل از تجزیه و تحلیل‌های آماری از کیفیت مطلوبی برخوردار بوده‌اند.

غلب در مجلات پزشکی مقالاتی به چشم می‌خورد که توسعه‌بدون کوچکترین اشاره‌ای به کیفیت داده‌های خام، یا استفاده از نمودارهای چند بعدی و رنگارانگ-که فقط زیبا هستند، و بهره‌گیری از آزمونهای پیشرفتی و پیچیده آماری و یا تکیه بر لغاتی نظری «معنی دار» که در اصل معنی دار از نظر آماری است نه از نظرات دیگر، و یا ارائه مقادیر P بسیار کوچک و در حد چند صدهزار بعنوان دلیلی جهت اثبات ادعای در حالیکه P احتمالات بوده و چیزی را به اثبات تعییر نمی‌سازد، نتایج را به نظر خوانندگان می‌رسانند.

در حالیکه آزمونهای آماری که هدف استفاده از آنها روشنگری و کمک به نتیجه‌گیری-نه توجیه ارتباطات، می‌باشد در تغییر ماهیت و کیفیت داده‌های خامی که در قالب قرمونهای آماری ریخته شده‌اند، کوچکترین نقشی را ایقا تمی نمایند. اصولاً پژوهشی را می‌توان صحیح یا بهتر از پژوهشی دیگر دانست و بر یافته‌های آن تکیه نمود که

مطالعه توسط پژوهشک یا پرسشنگر دیگر، مورد ارزیابی قرار داده و میزان ثبات داده‌های خام را بصورت کمی اندازه‌گیری نمود.

اندازه‌گیری میزان توافق

یکی از روش‌های متداول تعیین ثبات داده‌ها، اندازه‌گیری میزان توافق Agreement یا همخوانی Concordance است. برای محاسبه میزان توافق، می‌توان از روش‌های مختلفی بسته به نوع متغیر مورد مطالعه استفاده نمود. یکی از روش‌های محاسبه میزان توافق در مورد متغیرهای کیفی که در پژوهشکی کاربرد زیادی نیز دارد، محاسبه آماره کاپا Kappa Statistic می‌باشد که روش محاسبه آن بسیار ساده بوده و در قالب مثالی ذیلاً شرح داده خواهد شد.

روش محاسبه کاپا

فرض کنید دو نفر متخصص چشم پژوهش با تجربه ۵۰ نفر را مورد معاينه قرار می‌دهند و هدف آنها این باشد که مشخص نمایند که آیا این افراد چشمانشان سالم است یا غیرسالم. در جدول شماره ۱ تعداد و درصد یافته‌های این دو چشم پژوهشک آورده شده است.

جدول شماره ۱- میزان توافق مشاهده شده بین دو نفر چشم پژوهش *

		درصد افراد		تعداد افراد		
جمع		تشخیص چشم پژوهش A سالم	تشخیص چشم پژوهش A غیرسالم	جمع	تشخیص چشم پژوهش B سالم	تشخیص چشم پژوهش B غیرسالم
(p ₁)	۰/۴	(b) ۰/۸	(a) ۰/۳۲	(e+f) ۲۰	(f) ۴	(e) ۱۶
	۰/۶	(d) ۰/۵۶	(c) ۰/۴	(g+f) ۳۰	(h) ۲۸	(g) ۲
1.00	۱	۰/۶۴	۰/۳۶	۵۰	۳۲	۱۸
		(q _۱)	(p _۱)	(N)	(f+h)	(e+g)

اتفاق افتاده باشد. بنابراین نمی‌توان تنها بر مبنای میزان توافق مشاهده شده (۸۸ درصد) قضاوت نمود و ادعای کرد که این دو نفر چشم پژوهش در تشخیص موارد سالم و غیرسالم همخوانی بالائی با یکدیگر دارند. بلکه باید این سوال را مطرح نمود که «توافق ناشی از شанс بین این دو چشم پژوهش چقدر است؟» بنابراین برای پاسخ به سوال فوق ابتدا باید میزان توافق مورد انتظار ناشی از شанс را با روش زیر محاسبه نمود.

اصلًا در هر نوع اندازه‌گیری، مقداری خطای ذاتی وجود دارد. گرچه ممکن است نتوانیم در همه موارد خطاهای اندازه‌گیری را بطور کامل از بین ببریم، اما آگاهی از منابع خطای توجه به اثر و نقش خطاهای در نتایج مطالعه و آشنایی و استفاده از روشهایی که برای کاهش خطاهای بکار می‌روند، موجب می‌گردد که کیفیت نتایج مطالعه بهبود یافته و نتیجه‌گیری از تحقیق با واقعیت نزدیکتر گردد.

از آنجاییکه در اکثر موارد امکان تعیین مقدار حقیقی یک متغیر برای پژوهشگر وجود ندارد، بنابراین قابلیت تکرار یا ثبات داده‌ها که امکان اندازه‌گیری آن همواره وجود دارد، اهمیت بیشتری پیدا می‌کند. در تحقیقاتی که داده‌ها در چند مرحله یا با استفاده از چند روش مختلف یا توسط چند نفر جمع‌آوری می‌شود، محقق باید با عدد و ارقام نشان بدهد که داده‌های جمع‌آوری شده دارای ثبات قابل قبولی است.

ثبات دائم‌های جمع‌آوری شده با استفاده از یک روش جمع‌آوری مثل مصاحبه یا معاينه بالینی را می‌توان با تکرار اقلام مهم و کلیدی آن یا اندازه‌گیری اقلام کلیدی با چند روش مختلف، یا معاينه یا مصاحبه در صدی از افراد مورد

همانظریکه در جدول شماره ۱ دیده می‌شود، این دو چشم پژوهش A و B در مورد غیرسالم بودن ۱۶ نفر (۳۲ درصد افراد) و سالم بودن ۲۸ نفر (۵۶ درصد افراد) با یکدیگر توافق دارند. بنابراین مشاهده می‌شود که میزان توافق کلی این دو چشم پژوهش برابر است با $\frac{16+28}{88} = \frac{44}{88} = 0.50$ ، یا $a+d = 0/56 + 0/32 = 0/88$. این میزان توافق (۸۸ درصد) ظاهراً توافق خوبی بنظر می‌رسد. اما واقعیت اینست که در صدی از این توافق دیده شده ممکنست در اثر شанс

این اندکس یعنی عدد $75/0$ را «آماره کاپا» Kappa Statistic می‌نامند که با حرف «K» نشان داده می‌شود. بنابراین با توجه به روشی که در بالا برای محاسبه کاپا بکار بردیم می‌توان فرمول محاسبه کاپا را نوشت:

$$\frac{(\text{میزان توافق موردنظر ناشی از شناس}) - (\text{میزان توافق مشاهده شده})}{(\text{میزان توافق موردنظر ناشی از شناس}) - 1} = \text{آماره کاپا}$$

مطلوب داخل کادر «روش محاسبه آماره کاپا»

مقدار اندکس یا آماره کاپا می‌تواند بین $1+0$ (شامل صفر) تا عدد منفی باشد. بنابراین اگر حداقل میزان توافق (پس از در نظر گرفتن توافق شناسی) یا توافق کامل وجود داشته باشد، مقدار کاپا برابر با $+1$ خواهد بود. اگر توافق دیده شده با میزان توافق ناشی از شناس برابر باشد مقدار کاپا برابر با صفر و اگر میزان توافق دیده شده کمتر از میزان توافق ناشی از شناس باشد مقدار کاپا منفی خواهد بود. با توجه به اینکه مقدار کاپا عموماً بین صفر و $1-0$ باشد می‌توان بطور قراردادی میزان توافق را بشرح زیر بتبه نمود.

میزان توافق	مقدار کاپا
ضعیف	کمتر از $4/0$
متوسط تا خوب	$4/0-0/75$
خیلی خوب تا عالی	بیشتر از $75/0$

روش محاسبه آماره کاپا

$$\text{فرمول شماره ۱: } K = \frac{P_{\text{o}} - P_{\text{c}}}{1 - P_{\text{c}}}$$

P_{o} : میزان توافق مشاهده شده (در مثال فوق: $88/0$)

P_{c} : میزان توافق موردنظر ناشی از شناس که برابر است با:

$$P_{\text{c}} = (P_{\text{d}} \times P_{\text{v}}) + (q_{\text{d}} \times q_{\text{v}})$$

مقدار P_{c} در مثال فوق برابر است با:

$$P_{\text{c}} = (0/4 \times 0/36) + (0/6 \times 0/64) = 0/528$$

بنابراین مقدار آماره کاپا (K) در مثال فوق برابر است با:

$$K = \frac{0/88 - 0/528}{1 - 0/528} = 0/75$$

$$\text{فرمول شماره ۲: } K = \frac{2(ad - bc)}{P_{\text{d}}q_{\text{v}} + P_{\text{v}}q_{\text{d}}}$$

در مورد مثال فوق خواهیم داشت:

$$K = \frac{2(0/32 \times 0/56 - 0/8 \times 0/4)}{(0/4 \times 0/64) + (0/36 \times 0/6)} = 0/75$$

۱- چشم پزشک A تعداد 18 نفر را بعنوان «غیرسالم» تشخیص داده است که نسبت این تعداد به کل افراد معاینه شده (50 نفر) برابر با $36/0$ است. $(0/36 = 18/0)$.

۲- در مورد چشم پزشک B نسبت تشخیص‌های «غیرسالم» به کل افراد معاینه شده $4/0$ است. $(0/4 = 20/0)$. اگر تشخیص‌های «سالم» و «غیرسالم» توسط این دو نفر چشم پزشک بصورت «راندوم» توزیع شده باشد، انتظار داریم که توافق ناشی از شناس بین این دو نفر چشم پزشک در مسورد، «غیرسالم بودن» چشم $144/0$ باشد، $(0/144 = 0/36 \times 0/4)$.

۳- در مورد تشخیص‌های «سالم» نیز میزان توافق موردنظر ناشی از شناس بین دو نفر چشم پزشک برابر است با $384/0$ $(0/384 = 0/64 \times 0/6 = 0/32 \times 0/5)$.

۴- میزان توافق موردنظر ناشی از شناس در مورد هر دو شخص (سالم و غیرسالم) بین دو نفر چشم پزشک برابر است با مجموع توافق‌های شناسی فوق، یعنی $528/0$ $(0/528 = 0/144 + 0/384)$.

حالا با در نظر گرفتن میزان توافق ناشی از شناس یعنی $528/0$ درصد ($0/528$) می‌توان در مورد میزان توافق مشاهده شده که $88/0$ درصد ($0/88$) بود قضابت نمود. تفاوت این دو میزان توافق (توافق مشاهده شده و توافق شناسی) برابر است با $352/0$ درصد ($0/352 = 0/88 - 0/528$). بعبارت دیگر از توافق مشاهده شده $88/0$ درصد، $352/0$ درصد آن غیرشناختی است.

در اینجا این سوال مطرح می‌شود که: «مازاد توافق ناشی از شناس یا میزان توافق غیرشناختی، باید چقدر باشد تا بتوان میزان توافق مشاهده شده را ضعیف یا متوسط یا خوب یا عالی دانست؟». بعبارت دیگر، چگونه می‌توان مبنای عددی تبیه نمود که بر مبنای آن بتوان میزان توافق را رتبه‌بندی نمود.

برای این منظور، توافق کامل یعنی توافق 100 درصد ($1/00$) را مبنی قرار داده و میزان توافق ناشی از شناس را از آن کسر می‌کنیم. در مثال بالا میزان توافق ناشی از شناس $528/0$ بود که آنرا از توافق کامل یعنی عدد یک کسر می‌کنیم. در نتیجه خواهیم داشت: $0/472 = 0/528 - 0/0528$.

از طرقی تفاوت بین توافق دیده شده با توافق ناشی از شناس را تیز که قبل محاسبه کرده بودیم و برابر بود با $252/0$. حالا تبیه این دو عدد، یعنی $352/0$ و $252/0$ را محاسبه می‌کنیم. در نتیجه خواهیم داشت: $0/352 = 0/75 - 0/252$.

علل عدم توافق

نظر گرفتن نمایه‌های اعتبار اندازه‌گیری کرد. اغلب دیده می‌شد که تکنولوژی جدیدی را با تبلیغات گسترشده در امور بالینی وارد نموده و از آن بعنوان پیشرفته چشمگیر نام می‌برند. در حالیکه بعد از مدت کوتاهی نتایج تحقیقات نشان می‌دهد که تکنولوژی مزبور کمترین تغییری در مشکل یا بیماری مورد نظر ایجاد ننموده است. بعنوان مثال استفاده از ماموگرافی بعنوان تکنولوژی جدیدی برای تشخیص زودرس سرطان پستان سالها امید زیادی را برای کاهش مرگ و میر ناشی از سرطان پستان در جامعه پزشکی ایجاد نمود. اما تحقیقات انجام شده در زمینه غربالگری سرطان پستان در ادبینبورگ نشان داد که استفاده از ماموگرافی توانسته است امیدهای فوق را برآورده سازد(۲). با اینکه ماموگرافی بعنوان وسیله‌ای کارا جهت تشخیص زودرس سرطان پستان معرفی شده است ولی تاکنون میزان هم خوانی آن کمتر مورد بررسی قرار گرفته و عدم موفقتی کامل برنامه‌های غربالگری سرطان پستان با استفاده از ماموگرافی در کاهش میزان مرگ و میر ناشی از سرطان پستان ممکن است ناشی از محدودیتهای تکنولوژی ماموگرافی باشد که با بررسی‌های توافق و هم خوانی می‌توان به وجود آن پی برد.

از دیگر موارد مصرف تعیین میزان توافق و اندازه‌گیری آماره کاپا می‌توان به مقایسه تستهای تشخیصی یا غربالگری که از نظر سهولت انجام و یا مسائل اقتصادی با یکدیگر متفاوت می‌باشند اشاره نمود. بعنوان مثال بررسی میزان هم خوانی آزمون شکنندگی اسمزی Osmotic Fragility Test (O.F.T) (عنوان یک تست غربالگری، با روش تعیین شاخصهای خونی در تشخیص بیماری تالاسمی میتوان نشان داده است که با توجه به سهولت انجام و مخارج آزمون شکنندگی اسمزی، این تست جانشین مناسبی برای سایر تستهای غربالگری تالاسمی می‌باشد(۱).

استفاده از پرسشنامه بعنوان ابزاری جهت جمع‌آوری داده‌ها در اکثر تحقیقات اپیدمیولوژی، بالینی، اجتماعی متداول است. مهمترین مسئله در مورد پرسشنامه تعیین اعتبار درونی و پایایی آنست که می‌توان با استفاده از محاسبه میزان توافق به آن دست یافت. بعنوان مثال برای انجام بررسی جهانی باروری در کشورهای در حال توسعه که هدف آن تعیین عوامل مرتبط با باروری بود، قبل از انجام بررسی، پرسشنامه تهیه شده را دو بار و به فاصله یک تا دو هفته روی ۳۷۱ زن در Fiji مورد پیش آزمون قرار داده و

در مطالعاتی که افراد مورد مطالعه توسط دو یا چند نفر مورد معاينه قرار می‌گيرند میزان توافق آنها در مورد تشخيص، مرحله یا طبقه‌بندی بیماری می‌بايست روی نمونه‌ای از افراد مورد مطالعه و قبل از شروع به جمع‌آوری داده‌های مورد نیاز تحقیق انجام گیرد، و با تعیین میزان توافق بین معاينه کنندگان در صورت نیاز، در مورد بالا بردن میزان هم خوانی تصمیمات لازم در مورد استفاده از طرق مناسب (مثلاً آموزش مجدد) گرفته شود.

البته قبل از شروع مطالعه، می‌بايست روش معاينه (اندازه‌گیری) و طبقه‌بندی مورد بحث قرار گيرد تا همه معاينه کنندگان از اصول و روش واحدی برای تشخيص و طبقه‌بندی استفاده نمایند.

تعاریف واژه‌ها مثل، نرمال، غیرنرمال، بیمار، سالم، فعل، غیرفعال، سیگاری، غیرسیگاری و غیره، برحسب نوع بیماری و یا فاکتورهای مورد بررسی می‌بايست ضمن بحث و تبادل نظر با افراد مسئول جمع‌آوری داده‌ها مشخص شده و مورد قبول آنها قرار گیرد. پس از این مراحل میزان هم خوانی یا توافق در مورد معاينه کنندگان یا مشاهده کنندگان یا مصاحبه کنندگان محاسبه شده و در صورتیکه میزان توافق متناسبی بین آنها وجود نداشته و سوگیری وجود داشته باشد در مورد تکرار دوره آموزشی تصمیم‌گیری شود.

البته در مواردی که معاينه کنندگان در تخصص خود از تجربه زیادی هم برخوردار هستند، باز هم ممکنست که بین آنها توافق کاملی دیده نشود. علت این عدم توافق را می‌توان به «مشکل و پیچیده بودن» تشخیص نسبت داد که خود می‌تواند مربوط به وسائل کار و تجهیزات باشد. در واقع در چنین مواردی، میزان توافق محاسبه شده مربوط به تجهیزات و وسائل تشخیصی است نه افراد معاينه کننده.

در مواردی که از دستگاه جدیدی برای کمک به تشخیص استفاده می‌شود، اگر میزان توافق پایین باشد نایاب استفاده از آن دستگاه را قبل از آنکه تغییرات و یا تنظیم لازم در آنها صورت گرفته باشد، برای امور بالینی روزمره در درمانگاه یا بیمارستان توصیه نمود.

اصلًا قبل از استفاده از یک تکنولوژی جدید برای غربالگری، تشخیص و یا طبقه‌بندی جدید بیماری توصیه شده است که بعنوان اولین قدم میزان توافق آنرا مورد مطالعه قرار داده و در صورت بالابودن میزان توافق در تشخیص مراحل مختلف بیماری، در مرحله بعد باید کارایی آنرا با در

(نرمال و غیرنرمال) بالا می‌رود. این بالا رفتن میزان توان اتفاق کلی، حتی در صورت پایین بودن میزان توافق در امور در تشخیص‌های غیر نرمال نیز اتفاق می‌افتد.

همانطور که در جدول شماره ۱ دیده می‌شود، دو نفر چشم پزشک A و B، در مورد ۱۶ نفر از افراد مورد معاینه توافق دارند که این افراد چشم سالمی نداشته و نتیجه تشخیص خود را بعنوان «غیرنرمال» گزارش نموده‌اند. در مورد ۶ نفر نیز هر یک از چشم پزشکان تشخیص متفاوت با دیگری داده است. بنابراین در صد توافق در مورد تشخیص غیرنرمال برآبراست با $\frac{1}{73} = 0.0136$ یعنی $P = e/(e+f+g)$. این میزان توافق در مورد تشخیص‌های غیرنرمال از یک نمونه‌ای از جامعه بدست آمده است، بنابراین برای برآورده میزان توافق در مورد کل جامعه باید حدود اطمینان ۹۵

میزان توانق را با استفاده از فرمول زیر محاسبه نمود.

$$S.E. = \frac{95\% \text{ Confidence Interval}}{\text{Standard Error}}$$

S. E. = $\sqrt{\frac{P(1-P)}{n}}$ n = e + f + g
 نتیجه در مورد مثال فوق خطای معیار برابر است با:

$$S.E. = \sqrt{\frac{o/V \times o/V}{22}} = o/o$$

and will be held at the club

پس از محاسبه خطای معیار می توان با استفاده از فرمول زیر حدود اطمینان ۹۵ درصد میزان توافق را محاسبه نمود.

$$95\% \text{ CI} = P \pm 1/96 \times S.E.$$

رد مثال فوق، حدود اطمینان ۹۵ درصد
مورد تشخیص غیرنرمال بین دو نفر چشم
هد بود با:

حد بائنة اطمئنان ٢٠٢٤-٢٠٢٥

٩١ / ١٧٦ = ° + ° / ٧٣

95% CI = 0.05 to 0.91

بعبارت دیگر می توان با ۹۵ درصد اطمینان ادعا نمود که میزان توافق این دو نفر چشم پزشک در مورد تشخیص خیرنامال در جامعه بین ۵۵ درصد تا ۹۱ درصد می باشد.
با توجه به اینکه فاصله بین ۵۵ درصد توافق و ۹۱ درصد توافق زیاد می باشد بنابراین نمی توان بر مبنای کاپای ۰/۷۳ میزان توافق بین دو نفر چشم پزشک را در

میزان ترافق جوابهای هر سوال در دو مصاحبه را محاسبه نمودند و پس از اصلاح سوالاتی که پایانی بالائی نداشتند پرسشنامه جدیدی تهیه و بررسی را انجام دادند.^(۳)

تعیین میزان «توافق خوب» معمولاً بصورت اختیاری صورت می‌گیرد. اما با استفاده از حدود اطمینان ۹۵ درصد توافق نیز می‌توان «توافق خوب» را تعیین نمود. بدین ترتیب که اگر حجم نمونه مورد مطالعه کافی بوده و حدود اطمینان ۹۵ درصد توافق دربرگیرنده $50/0$ (۵۰ درصد) باشد می‌توان میزان توافق را «ضعیف» دانست. بعنوان مثال: اگر حدود اطمینان ۹۵ درصد کاپای توافق $80/0$ تا $85/0$ باشد، با اینکه حد بالائی توافق 85 درصد است نباید این توافق را عالی یا خوب دانست، زیرا حد پائینی آن $30/0$ (۳۰ درصد) بوده و حدود اطمینان دربرگیرنده توافق 50 درصد می‌باشد.

حدود توافق:

کاپای محاسبه شده از یک نمونه بدست آمده است. چه بسا اگر نمونه دیگری از بیماران توسط همین دو نفر چشم پژشک مورد معاینه قرار گیرند و مقدار کاپا برای توافق مجدداً محاسبه گردد مقدار آن با کاپای محاسبه شده قبلی تفاوت داشته باشد. بنابراین همانطوریکه برای یک میانگین یا درصد حدود اطمینان محاسبه می‌شود در مورد کاپا نیز می‌توان بعنوان مثال ۹۵ درصد حدود توافق را محاسبه نمود تا بتوان ادعا نمود که توافق واقعی (در مورد جامعه) با ۹۵ درصد اطمینان بین دو محدوده محاسبه شده قرار گرفته است.

بر حسب موضوع تحقیق یا بیماری مورد مطالعه ممکنست لازم باشد که میزان توافق دو نفر معاینه کننده را برای تشخیص های نرمال و تشخیص های غیرنرمال جداگانه محاسبه نمود. زیرا هدف اصلی از انجام تست یا تشخیص ممکنست بیشتر جداسازی افراد نرمال باشد و در مورد بیماری یا تستی دیگر هدف اصلی بیشتر جداسازی افراد غیرنرمال باشد.

در مواردی هم هر دو حالت یعنی نرم‌مال بودن یا غیرنرم‌مال بودن اهمیت دارد و باید میزان توافق را برای هر دو حالت محاسبه ننمود.

دلیل دیگر سرای محاسبه جدأگانه میزان توافق تشخیص‌های نرمال و غیرنرمال اینست که در موقعیکه درصد افراد غیرنرمال در نمونه مورد بررسی کوچک باشد، به علت بالا بودن درصد افراد نرمال، میزان توافق کلی

آن ۹۵٪ تا ۶۹٪ است.
در ضمن اگر مقادیر e + g + f کوچکتر از ۵ باشد فرمول
محاسبه خطای معیار معتبر نبوده و برای محاسبه ۹۵ درصد
حدود اطمینان باید از روش دیگری استفاده نمود.

تشخیص‌های غیرنرمال توافق قابل توجهی قلمداد نمود.
اگر با روش فوق، میزان توافق برای تشخیص‌های نرمال
و حدود اطمینان ۹۵ درصد آنرا محاسبه نمائیم به این نتیجه
می‌رسیم که میزان تعلق ۸۲٪ و حدود اطمینان ۹۵ درصد

منابع:

2. Roberts MM, et al. Edinburgh Trial of Screening for Breast Cancer: Mortality at Seven years. *Lancet* 1990; 335: 241-6.
3. O'Muircheartaigh CA, Marckwardt AM. An Assessment of the Reliability of World Fertility Survey Data. London: World Fertility Survey Conference, 1980.

Keywords: Models, Statistical/ Research/ Statistics/ Statistic and Numerical Data