

تعیین حجم نمونه به کمک G*Power

امین رضا امینی

خرداد ۱۴۰۱

رئوس مطالب

- اهمیت حجم نمونه
- عوامل تاثیر گذار بر حجم نمونه
- مطالعات توصیفی و تحلیلی
- خطاهای نوع اول و دوم و توان آزمون
- اندازه اثر
- اطلاعات پیشین و مطالعه دم دستی (pilot study)
- ویژگی ها و اصول یک مطالعه دم دستی
- معرفی نرم افزار G*Power

رئوس مطالب

- حجم نمونه در مطالعات توصیفی
 - میانگین
 - نسبت
- حجم نمونه در مطالعات تحلیلی
 - آزمون تی یک نمونه ای
 - آزمون تی دو نمونه ای
 - زوجی
 - مستقل
 - آزمون مقایسه نسبت ها
 - تحلیل واریانس یک طرفه
 - آزمون تحلیل اندازه های تکراری
 - آزمون کای دو
 - آزمون همبستگی

مطالعات توصیفی و تحلیلی

مطالعات توصیفی

نمونه هایی از مطالعات توصیفی:

- ارزیابی تیتراژ آنتی بادی ضد آنتی ژن سطحی ویروس هپاتیت B در دانشجویان ورودی سال ۱۳۹۸ دانشگاه علوم پزشکی شیراز، ایران
- بررسی شیوع مایکوباکتریوم های غیرسلی در بیماران مراجعه کننده به آزمایشگاه رفرانس سل کشوری
- تعیین میانگین سطح سرب مایع مغزی- نخاعی کودکان مبتلا به تشنج ایدیوپاتیک

مطالعات توصیفی

- در مطالعات توصیفی صرفاً هیچ گونه رابطه سنجی بین متغیرها وجود ندارد.
- توصیفی از جامعه تحت مطالعه و وضعیت موجود یک مشکل یا بحران یا رویداد بدون توجه به عامل یا اثر آن
- برآورد میانگین، میانه، نسبت، شیوع و ...
- به آزمون های آماری نیاز نیست.

مطالعات توصیفی

- هدف ← تعیین میانگین یک متغیر کمی و یا نسبت یک متغیر کیفی
- در این مطالعات پژوهشگر به دنبال آن است که برآورد میانگین یا نسبت مورد نظر را با حداکثر فاصله معینی از واقعیت (دقت d) به دست آورد.
- میزان دقت d یا خطای قابل قبول یک مساله تخصصی است و مبتنی بر حوزه تحقیقاتی
- هرچه دقت بیشتر، حجم نمونه بیشتر

مطالعه توصیفی - برآورد میانگین

$$n = \frac{(Z_{1-\frac{\alpha}{2}})^2 \sigma^2}{d^2}$$

d = دقت

اگر محقق پیش فرضی در مورد میزان d مطلوب نداشته باشد
۱. از نامساوی زیر استفاده شود

$$0.1\sigma < d < 0.32\sigma$$

۲. با استفاده از مطالعه گذشته، می توان نصف فاصله اطمینان گزارش شده برای میانگین را به عنوان d در نظر گرفت

مطالعه توصیفی – برآورد میانگین

• محققى مایل است تا میانگین سطح سرب مایع مغزی- نخاعی کودکان مبتلا به تشنج ایدیوپاتیک

را در سطح خطای نوع اول ۰/۰۵ تعیین نماید. بر اساس مطالعه پیشین میانگین این سطح برابر با

$۳/۰۷ \pm ۳/۴۳$ شده است.

مطالعه توصیفی – برآورد نسبت

$$n = \frac{(Z_{1-\frac{\alpha}{2}})^2 pq}{d^2}$$

دقت = d

اگر محقق پیش فرضی در مورد میزان d مطلوب نداشته باشد

$$0.1p < d < 0.2p$$

مطالعه توصیفی – برآورد نسبت

• مثال)

محقق می خواهد شیوع مایکو باکتریوم غیرسلی از نوع چلونی را در بیماران مراجعه کننده به

آزمایشگاه رفرانس سل کشوری برآورد نماید ($\alpha = 0.05$). براساس مطالعات پیشینی ، شیوع این نوع

از مایکو باکتریوم حدوداً برابر با ۴۴ درصد است.

متناهی بودن یا نبودن حجم جامعه هدف

- آیا علاوه بر پراکندگی صفت مورد نظر در جامعه، حجم جامعه هدف که نتایج مطالعه به آن تعمیم خواهد یافت، در تعیین حجم نمونه تاثیری دارد؟

- حجم جامعه هدف > 10000 نفر  باید برای بدست آوردن حجم نمونه مناسب،

حجم جامعه را دخیل نمود.

$$n' = \frac{n}{1 + \frac{n}{N}}$$

مطالعات تحلیلی

- علاوه بر توصیف و بیان وضعیت موجود، در پی سنجش ارتباطات بین متغیرها و یافتن اختلافات بین گروه های تحت مطالعه هستیم.

- نیازمند به

- فرضیات H_0 و H_1

- آزمون های آماری

- محاسبه P-Value

- توان آزمون و حجم نمونه

- خطاهای آماری

- نوع اول

- نوع دوم

خطاهای آماری و توان آزمون

• خطای نوع اول

• خطای نوع دوم

• توان آزمون

Type I and Type II Error

Null hypothesis is ...	True	False
Rejected	Type I error False positive Probability = α	Correct decision True positive Probability = $1 - \beta$
Not rejected	Correct decision True negative Probability = $1 - \alpha$	Type II error False negative Probability = β

رابطه بین توان آزمون و حجم نمونه

- با افزایش حجم نمونه، توان آزمون (احتمال رد فرض صفر وقتی که نادرست است) افزایش می یابد.
- روال کار بدین صورت است:
 ۱. پیش از انجام مطالعه، خطای نوع اول را ثابت در نظر بگیریم.
 ۲. حجم نمونه را محاسبه نماییم.
 ۳. آزمون آماری مناسب را انتخاب نماییم.
 ۴. در صورت رد نشدن فرض صفر، توان آزمون را محاسبه نماییم.

اندازه اثر و معنی داری آماری

P-value

- میزان شانسی بودن اختلاف یا ارتباط مشاهده شده.
- تفاوت ها و یا ارتباطات مشاهده شده تا چه میزان می تواند صرفاً بر اثر شانس و تصادف است.
- کوچک بودن P-value نشاندهنده اثر کمتر تصادف است.
- تلویحا بیان می کند که ارتباطات و یا اختلافات مشاهده شده در بین متغیرها ذاتی است و نه براساس شانس.

آنچه که P-value به ما نمی گوید...

مثال : در مطالعه ای رابطه بین مصرف سیگار و ابتلا به سرطان ریه و معده معنی دار شده است ($P < 0.05$).

- آیا می توان دریافت که سیگار کشیدن به چه میزان احتمال بروز سرطان ریه و معده را افزایش می دهد؟
- کدام یک از این دو سرطان با شدت بیشتری با سیگار کشیدن مرتبط است؟
- در مطالعات تحلیلی تنها مشاهده یک ارتباط معنی دار از نظر آماری کافی نیست و باید میزان این ارتباط نیز تعیین گردد.

اندازه اثر

- کمی کردن میزان تاثیر مواجهه (exposure)، مداخله (intervention)، تیمار (treatment) و یا اثر (effect) و یا کمی کردن شدت رابطه (relation or association)
- شدت ارتباط - شدت اختلاف
- متاآنالیز

انواع اندازه اثر ها

- نسبت شانس OR
- خطر نسبی RR
- ضریب همبستگی پیرسون ρ
- d کوهن
- و

اندازه اثر و حجم نمونه

- هرچه اندازه اثر بزرگتر باشد حجم نمونه مورد نیاز کوچکتر خواهد بود. مثال؟
- برآوردی بهینه و مناسب از اندازه اثر نکته کلیدی در تعیین حجم نمونه است.
- اینکه به مقداری منطقی و موجه از اندازه اثر دست یابیم همواره کاری آسان نیست.
- بهترین برآورد از اندازه اثر را می توان از روی مطالعات گذشته و یا مطالعه دم دستی (pilot study) یافت.

انواع اندازه اثرها

Cohen, regarded as the deity of power analysis, (1977, 1988) justifies these levels of effect sizes.

	Effect size Index	Small	Medium	Large
t-test on Means	d	0.20	0.50	0.80
t-test on Correlations	r	0.10	0.30	0.50
F-test ANOVA	f	0.10	0.25	0.40
F-test regression	f ²	0.02	0.15	0.35
Chi-Square Test	w	0.10	0.30	0.50

مطالعه دمِ دستی (Pilot Study)

- مطالعات دمِ دستی ، مطالعاتی در مقیاس کوچک هستند که هدف آن بررسی این است که آیا اجزای حیاتی یک مطالعه اصلی امکان‌پذیر است یا خیر.
- پاسخ به این سوال : "آیا مطالعه اصلی می‌تواند به روشی که برنامه ریزی شده است انجام شود یا باید برخی از مؤلفه‌ها را تغییر داد؟"

• مثال

- تعیین حجم نمونه مناسب برای پروژه اصلی
- جلوگیری از اتلاف وقت و منابع

مطالعه دمِ دستی

- اندازه نمونه ممکن است در مطالعات آزمایشی متفاوت باشد (مقالات مختلف محاسبات اندازه نمونه متفاوتی را ارائه می دهند)
- جامعه مطالعه دم دستی که نمونه از آن تشکیل شده است باید با مطالعه اصلی یکسان باشد.
- شرکت کنندگان در مطالعه دم دستی نباید وارد مطالعه اصلی شوند. به این دلیل که شرکت کنندگان ممکن است رفتار بعدی خود را تغییر دهند.

مطالعه دم دستی

Flat Pilot Study Sample Size Rules of Thumb for Two-Group Studies

Reference	Recommended Pilot Study Sample Size	Comments
Birkett and Day (1994)	20	Suggested 20 for Internal pilot studies.
Browne (1995)	30	Mentions that the use of 30 is commonplace at the time
Kieser and Wassmer (1996)	20 to 40	Use when main trials are between 80 and 250 and using UCL.
Julious (2005)	24	Recommended minimum of 12 subjects per group.
Sim and Lewis (2011)	≥55	Use for small to medium effect sizes to minimize combined size.
Teare, et al. (2014)	≥70	Based on an extensive simulation study.

Note: These rules of thumb for two groups can be adapted for single-group or multi-group studies by multiplying the recommended pilot study sample size by an appropriate adjustment factor (e.g., 0.5 for a study with only one group or 1.5 for a study with three groups, etc.).

دقت و اندازه اثر

- در مطالعات توصیفی ← دقت
- در مطالعات تحلیلی ← اندازه اثر

G*Power

- <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower>

آزمون تی یک نمونه ای

$$H_0 : \mu - \mu_0 = 0$$

$$H_1 : \mu - \mu_0 \neq 0.$$

• فرضیات صفر و جانشین

• در این گونه از آزمون های یک نمونه ای، معمولا مقایسه نمونه ای از جمعیت با شاخص های

استاندارد یا آمارهای جمعیتی بسیار کاربرد دارد.

$$d = \frac{\mu - \mu_0}{\sigma}$$

• در چه میزان پراکندگی اختلاف بین گروه ها مشاهده شده است؟

• small $d = 0.2$

• medium $d = 0.5$

• large $d = 0.8$

• استاندارد سازی اختلافات

آزمون تی یک نمونه ای

مثال:

- متوسط قد مردان بزرگسال ایرانی ۱۷۴ سانتی متر تخمین زده شده است .
- محقق ادعا دارد که همبستگی مثبتی میان قد و هوش افراد وجود دارد. اگر چنین باشد انتظار بر این است که متوسط قد دانشجویان پسر دانشگاه صنعتی شریف به صورت کلی بزرگتر از متوسط قد مردان ایرانی باشد. این محقق به منظور آزمودن ادعای خود، چه تعداد نمونه ی تصادفی از دانشجویان پسر این دانشگاه باید بردارد؟ (برآورد دم دستی از میانگین و انحراف معیار قد دانشجویان این دانشگاه به ترتیب برابر با ۱۷۶ و ۴/۵ است).

آزمون تی یک نمونه ای

(مثال)

- برآورد سن افراد در شروع به مصرف سیگار در شهرستان نیشابور

آزمون تی دو نمونه ای

زوجی

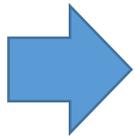
مستقل

آزمون تی زوجی (وابسته)

دو نمونه زوج شده داریم

$$H_0: \mu_y = \mu_x$$

$$H_1: \mu_y \neq \mu_x$$



$$H_0: \mu_z = 0$$

$$H_1: \mu_z \neq 0$$

X_i ← نمونه ۱ (یا قبل)

Y_i ← نمونه ۲ (یا بعد)

$$z_i = y_i - x_i$$

$$d_z = \frac{|\mu_z|}{\sigma_z} = \frac{|\mu_x - \mu_y|}{\sqrt{\sigma_x^2 + \sigma_y^2 - 2\rho_{xy} \sigma_x \sigma_y}}$$

آزمون تی زوجی (وابسته)

مثال ۱) اپلیکیشنی برنامه کاهش وزن هشت هفته ای را به بازار عرضه کرده و ادعا دارد که در پایان برنامه، هر فرد به طور متوسط، ۵ کیلو وزن کم می کند.

شما برنامه را مطالعه کرده اید و معتقدید که برنامه آنها از نظر علمی نادرست است و نمی تواند موثر باشد. با بودجه محدودی که در اختیار دارید، می خواهید این فرضیه را بیازمایید که این برنامه به کاهش وزن افراد کمکی نمی کند.

راهبرد شما بدین گونه است که یک نمونه تصادفی از افراد به دست آورید و آنها را در برنامه قرار دهید. وزن آنها را در ابتدای برنامه اندازه گیری کنید و سپس در پایان برنامه دوباره وزن آنها را اندازه بگیرید. بر اساس برخی تحقیقات

قبلی، شما معتقدید که انحراف معیار اختلاف وزن در طول هشت هفته ۷ کیلوگرم خواهد بود. اکنون می خواهید

بدانید که برای آزمایش فرضیه خود باید چند نفر را در برنامه ثبت نام کنید.

آزمون تی زوجی (وابسته)

مثال (۲)

محقق می خواهد تفاوت بین دست غالب و دست غیر غالب را از نظر مهارت و چابکی بررسی کند. او آزمایشی را طراحی می کند که در آن هر آزمودنی ۱۰ مهره کوچک را روی میز در یک کاسه قرار می دهد، یک بار با دست غالب و یک بار با دست غیر غالب. او زمان مورد نیاز (برحسب ثانیه) در هر دور را برای تکمیل کار اندازه می گیرد. او انتظار دارد که میانگین اختلاف زمان ۳ ثانیه باشد و دست غالب با انحراف معیار ۱۰ کارآمدتر باشد. این محقق به چه حجمی از نمونه نیازمند است؟

آزمون تی مستقل

آزمون تی مستقل

مثال

یک متخصص تغذیه بالینی می خواهد دو رژیم غذایی متفاوت A و B را برای بیماران دیابتی تجویز و نتایج برآمده از دو گروه را مقایسه نماید. او ادعا می کند که رژیم A (گروه ۱) بهتر از رژیم B (گروه ۲) از نظر پایین تر بودن قند خون خواهد بود. او قصد دارد یک نمونه تصادفی از بیماران دیابتی بگیرد و آنها را به طور تصادفی به یکی از این دو رژیم غذایی اختصاص دهد. در پایان آزمایش که ۶ هفته طول می کشد، آزمایش قند خون ناشتا برای هر بیمار انجام می شود.

او همچنین انتظار دارد که میانگین تفاوت در اندازه گیری گلوکز خون بین دو گروه حدود ۱۰ میلی گرم در دسی لیتر باشد. علاوه بر این، او همچنین انحراف معیار توزیع گلوکز خون برای رژیم A را ۱۵ و انحراف معیار برای رژیم B را ۱۷ فرض می کند. این متخصص تغذیه در هر گروه به چه تعداد بیمار-با فرض گروه های مساوی- نیاز دارد؟

در اینجا تا زمانی که تفاوت بین آنها مقدار طبق فرض مساله به درستی رعایت شود اهمیتی ندارد که میانگین گروه های ۱ و ۲ دقیقا به چه مقدار باشند. در اینجا تفاوت مد نظر برابر با ۱۰ است. بنابراین میانگین های دو گروه را به صورت های

۰ و ۱۰،

۵ و ۱۵،

-۲ و ۸

و ...

همه قابل قبول خواهند بود.

آزمون تی مستقل – ملاحظات

- An important technical assumption is the normality assumption. If the distribution is skewed, then a small sample size may not have the power shown in the results, because the value in the results is calculated using the method based on the normality assumption. We have seen that in order to compute the power or the sample size, we have to make a number of assumptions. These assumptions are used not only for the purpose of calculation, but are also used in the actual t-test itself. So one important side benefit of performing power analysis is to help us to better understand our designs and our hypotheses.
- We have seen in the power calculation process that what matters in the two-independent sample t-test is the difference in the means and the standard deviations for the two groups. This leads to the concept of effect size. In this case, the effect size will be the difference in means over the pooled standard deviation. The larger the effect size, the larger the power for a given sample size. Or, the larger the effect size, the smaller sample size needed to achieve the same power. So, a good estimate of effect size is the key to a good power analysis. However, it is not always an easy task to determine the effect size. Good estimates of effect size come from the existing literature or from pilot studies.

آزمون آنالیز واریانس یک طرفه

$$f = \sigma_m / \sigma$$

• اندازه اثر

• σ_m انحراف معیار میانگین های گروه ها است.

• σ انحراف معیار درون-گروهی است که برای همه گروه ها مقداری یکسان و ثابت در نظر گرفته

می شود.

Cohen (1969, p.348) defines the following effect size conventions:

- small $f = 0.10$
- medium $f = 0.25$
- large $f = 0.40$

آزمون آنالیز واریانس یک طرفه

SD σ within each group

6

نکته کاربردی برای پر کردن

- در اینجا باید مقدار σ را قرار دهیم و چون این پارامتر برای همه گروه ها یکسان در نظر گرفته می شود پس از انحراف معیار گروه ها میانگین می گیریم.

آزمون آنالیز واریانس یک طرفه

- Example: Mean pH and Mean Metabolite Ratios (\pm SD) in Normal Control Group, Astrocytoma Group, Lymphoma Group, Metastasis Group and Abscess Group

	Normal Control (n = 11) Mean \pm SD	Astrocytoma (n = 13) Mean \pm SD	Lymphoma (n = 4) Mean \pm SD	Metastasis (n = 5) Mean \pm SD	Abscess (n = 6) Mean \pm SD	<i>P</i> *
pH	7.07 \pm 0.10	7.15 \pm 0.20	7.45 \pm 0.32	7.48 \pm 0.56	7.56 \pm 0.64	0.058
PME/PDE	0.47 \pm 0.07	0.56 \pm 0.10	0.60 \pm 0.15	0.68 \pm 0.19	0.65 \pm 0.32	0.034
PME/Pi	1.60 \pm 0.19	1.83 \pm 0.50	1.56 \pm 0.61	1.69 \pm 0.35	2.00 \pm 1.01	0.304
PDE/Pi	3.44 \pm 0.49	3.27 \pm 0.62	2.57 \pm 0.60	2.76 \pm 1.27	3.18 \pm 0.78	0.039
PME/ATP	0.21 \pm 0.03	0.24 \pm 0.05	0.20 \pm 0.05	0.23 \pm 0.04	0.28 \pm 0.08	0.152
PDE/ATP	0.45 \pm 0.09	0.44 \pm 0.10	0.35 \pm 0.11	0.37 \pm 0.15	0.48 \pm 0.18	0.340
PME/PCr	0.34 \pm 0.07	0.43 \pm 0.11	0.30 \pm 0.08	0.36 \pm 0.10	0.42 \pm 0.09	0.028
PDE/PCr	0.75 \pm 0.20	0.77 \pm 0.18	0.51 \pm 0.16	0.59 \pm 0.28	0.73 \pm 0.27	0.049
PCr/ATP	0.61 \pm 0.07	0.58 \pm 0.15	0.69 \pm 0.12	0.66 \pm 0.11	0.67 \pm 0.13	0.314
PCr/Pi	4.74 \pm 0.74	4.36 \pm 0.95	5.51 \pm 2.52	4.72 \pm 0.48	4.86 \pm 2.18	0.640
ATP/Pi	7.76 \pm 0.64	7.90 \pm 2.51	8.37 \pm 4.34	7.28 \pm 1.39	7.61 \pm 4.12	1.000

آزمون آنالیز واریانس یک طرفه

حجم نمونه مطالعه گذشته؟

اندازه اثر pH؟

حجم نمونه کنونی؟

آزمون آنالیز واریانس یک طرفه

مثال (۲)

آراف فرکشنال یکی از دستگاه‌هایی است که به منظور زیبایی و رفع ایرادات پوست مورد استفاده قرار می‌گیرد. در این دستگاه از لیزر RF جهت جوانسازی پوست استفاده می‌شود و به شکلی مطلوب بر روی پوست اثرگذار است. به نحوی که پس از انجام آراف فرکشنال پوست می‌توان بازسازی و جوانسازی را به وضوح در ناحیه مورد نظر مشاهده کرد. این روش به عنوان یکی از شیوه‌های کم‌تهاجمی شناخته می‌شود که برای انواع پوست در سنین مختلف کاربرد دارد. جوانسازی دوباره با آراف فرکشنال توسط عملکرد مطلوب دستگاه با تاثیر بر روی لایه‌های سطحی و عمقی پوست انجام‌پذیر است تا نتیجه‌ای دلخواه حاصل گردد.

آزمون آنالیز واریانس یک طرفه

- محقق می خواهد با انجام یک کارآزمایی بالینی سه گروهه، پاسخ به درمان را (رفع شل و سست بودن پوست صورت و گردن) در سه سطح ۵۲-۵۷ درجه، ۶۲ درجه و ۶۷ درجه سانتی گراد پس از گذشت سه ماه بررسی نماید. حجم نمونه لازم؟
- مفروضات مساله:
- $\alpha=0.05$
- $\beta=0.2$
- response = laxity after 3 months

آزمون آنالیز واریانس یک طرفه

TABLE 6. Clinical Results

	<i>Arm 1 (Topical Only): 52°C–57°C, 3 sec; n = 5</i>				<i>Arm 2 (Standard Infiltration): 62°C, 3 sec; n = 5</i>				<i>Arm 3 (Standard Infiltration): 67°C, 3 sec; n = 5</i>			
	<i>Rhytides</i>		<i>Laxity</i>		<i>Rhytides</i>		<i>Laxity</i>		<i>Rhytides</i>		<i>Laxity</i>	
	<i>3 mo</i>	<i>6 mo</i>	<i>3 mo</i>	<i>6 mo</i>	<i>3 mo</i>	<i>6 mo</i>	<i>3 mo</i>	<i>6 mo</i>	<i>3 mo</i>	<i>6 mo</i>	<i>3 mo</i>	<i>6 mo</i>
Average grade improvement (SD)	0.47 (1.37)	0.87 (1.07)	0.27 (0.75)	0.50 (0.78)	1.13 (1.12)	1.13 (0.87)	0.73 (0.25)	0.97 (0.49)	1.87 (0.51)	2.27 (1.12)	0.97 (0.55)	1.07 (0.42)
Min	2.00	2.00	1.17	1.17	3.00	2.00	1.00	1.50	2.33	3.67	1.50	1.50
Max	-1.33	-0.67	-0.83	0.83	0.00	0.00	0.50	0.50	1.00	1.00	0.33	0.50
Responder rate	80%				100%				100%			
Average before score	3.47		2.20		3.93		2.70		5.73		3.10	
Percent improvement (SD)	13 (42)	25 (56)	12 (34)	25 (38)	29 (15)	29 (24)	27 (8)	36 (15)	33 (8)	40 (21)	31 (15)	34 (16)

آزمون برابری نسبت ها

مشخص شده است که نوع خاصی از ضایعات پوستی در صورت عدم درمان در ۳۰ درصد بیماران به سرطان تبدیل می شود. دارویی در بازار وجود دارد که احتمال ابتلا به سرطان را تا ۲۰ درصد کاهش می دهد. یک شرکت داروسازی در حال تولید داروی جدیدی برای درمان ضایعات پوستی است، اما این داروی جدید تنها زمانی ارزش دارد که احتمال ابتلا به سرطان را به ۱۵ درصد و یا بیشتر کاهش دهد.

این شرکت داروسازی در نظر دارد با بیمارانی که به صورت تصادفی در دو گروه کنترل (درمان نشده) و گروه درمان قرار گرفتند، مطالعه ای انجام دهد. محاسبه کنید که این شرکت با توان $0/8$ در آلفای برابر با $0/05$ به چه تعداد آزمودنی نیاز دارد.

- همه این کارهایی که انجام شده خوب و نیکو است، اما به نظر می رسد که آزمایش دو طرفه در این شرایط چندان منطقی نباشد چرا که ما می خواهیم برای دارویی آزمایش کنیم که احتمال سرطان را کاهش می دهد نه دارویی که احتمال را افزایش می دهد.
- در این مورد ممکن است بخواهیم از یک تست یک دامنه استفاده کنیم که با تغییر ورودی "tail(s)" به راحتی قابل تنظیم است.

آزمون کای دو – آزمون استقلال

- آزمونی برای متغیرهای کیفی (رسته ای)
- جدول توافقی
- فرض صفر: استقلال سطر و ستون ها
- اندازه اثر

$$\phi = \sqrt{\frac{\chi^2}{N}}$$

آزمون کای دو – آزمون استقلال

- پزشکی که در سانتر بیماران کرونایی مشغول به کار بوده، مشاهده کرده است که گروه خونی بیماران در میزان ابتلا به کرونا تاثیر دارد. وی برای سنجش ادعای خود به چه حجمی از نمونه نیاز دارد ($\alpha=0.05$ و $\beta=0.2$)؟

آزمون کای دو- آزمون استقلال

• براساس یک مطالعه دم دستی اطلاعات زیر به دست آمده است:

گروه خونی	مثبت	منفی
O	۹	۴
سایر	۵	۹
کل	۱۴	۱۳

آزمون همبستگی

در ادامه

- مطالب پیشرفته تر شامل
- تحلیل رگرسیون
- آزمون های ناپارامتری
- مدل های بقا (رگرسیون کاکس و ...)