



درسنامه استارت

آمار

تیم تالیف گروه مدمشاور

توجه: در کنار این درسنامه، اطلس تمام تصاویر با کیفیت بالا به صورت یک فایل pdf در اختیارتون قرار خواهد گرفت. برای دریافت این اطلس کد زیر را اسکن کنید و یا به کانال تلگرامی مدمشاور مراجعه کنید.



فهرست مطالب

۱۰	سوالات آزمون ها	۱	۱. متغیرها و نمودارهای توزیع فراوانی
۱۹	۳. احتمالات	۱	۱. کیفیت
		۱	۲. کمی
۱۹	قانون ضرب احتمالات	۱	نمودارها
۱۹	قانون جمع احتمالات	۱	۱. برای اطلاعات کیفی
۲۰	توزیع دو جمله ای	۲	۲. برای اطلاعات کمی
۲۰	سوالات آزمون ها	۳	سوالات آزمون ها
۲۵	نظری بیز (Bayes)		
۲۵	♦ حساسیت آزمون	۷	۲. شاخص های مرکزی و شاخص های پراکندگی
۲۵	♦ ویژگی آزمون		
۲۵	♦ کاربردها		
۲۶	روش تست های متعدد	۷	شاخص های مرکزی
۲۶	♦ متوالی Sequential (۱)	۷	شاخص های پراکندگی
۲۶	♦ موازی (۲)	۷	مرکزی ها
۲۶	♦ منفی کاذب	۷	۱. میانگین یا Mean
۲۶	♦ مثبت کاذب	۷	۲. میانه یا Median
۲۶	♦ ارزش اخباری	۸	۳. مد یا نما یا Mode
۲۶	راه حل تستی	۸	۴. صدک ها و دهک ها
۲۸	♦ Reliability (قابلیت اعتماد)	۸	شاخص های پراکندگی
۲۸	♦ Validity (قابلیت اعتبار)	۸	۱. دامنه تغییرات یا Range
۲۸	سوالات آزمون ها		۲. دامنه میان چارکی
		۸	(Interquartile Range)
۳۹	۴. توزیع احتمالات	۸	۳. میانگین انحرافات (Mean deviation)
		۹	۴. انحراف معیار
۳۹	توزیع پواسون	۹	۵. واریانس
۳۹	سوالات آزمون ها	۹	۶. ضریب تغییرات (CV)

خطای معیار میانگین یا خطای استاندارد (SEM)	۴۱	توزیع نرمال
۷۱	۴۲	نمونه سوالات تمرینی
۷۲ حجم نمونه	۴۴	توزیع‌های چوله شده (skewed)
۷۲ ♦ فرمول اول (برای صفت کمی)	۴۴	سوالات آزمون‌ها
♦ فرمول نسبت حجم (برای صفت کیفی، نسبت یا شیوع)		
۷۲	۵۵	۵. اصول تفسیر آماری
۷۳ حدود اطمینان میانگین جامعه		
۷۳ حدود اطمینان برای شیوع (نسبت)	۵۵	آزمون فرضیه
۷۴ سوالات آزمون‌ها	۵۵	♦ فرضیه صفر (H_0)
	۵۵	♦ فرضیه H_1
۷۹ ۸. انواع مطالعات	۵۵	انواع خطاها
۷۹ ♦ (۱) توصیفی	۵۵	توان آزمون یا Power
۷۹ ♦ (۲) تحلیلی یا آنالیتیک	۵۶	P - value
۷۹ مداخله دارد (بالینی‌ها)	۵۶	خلاصه تفاوت معنادارها
۷۹ مداخله ندارد:	۵۷	سوالات آزمون‌ها
۸۰ ♦ مطالعات توصیفی		
♦ مطالعه محیطی Ecologic یا تطبیقی		
۸۰ Correlation	۶۱	۶. آزمون کای اسکوئر و آزمون‌های معنادار
♦ مطالعه مقطعی یا مطالعه شیوع یا Cross sectional	۶۱	۱. اطلاعات کمی
۸۰ ♦ مطالعه مورد - شاهدی یا گذشته نگر یا - Case Control	۶۱	♦ الف) مقایسه میانگین دو گروه با هم
۸۰ ♦ کوهورت یا همگروهی یا آینده نگر یا مطالعه وقوع (بروز)	۶۱	♦ ب) مقایسه میانگین بیش از دو گروه
۸۱	۶۱	♦ ج) مقایسه میانگین دو جز وابسته
۸۲ انواع کوهورت	۶۱	۲. اطلاعات کیفی
۸۲ ♦ کوهورت گذشته نگر	۶۱	♦ الف) مستقل
۸۲ ♦ کوهورت ترکیبی	۶۲	♦ ب) غیر مستقل Mc Nemor
۸۲ خلاصه مبحث انواع مطالعات اولیه	۶۲	همبستگی بین صفات
مطالعات مداخله ای یا کارآزمایی بالینی (تجربی)	۶۲	♦ صفت مستقل
۸۳	۶۲	♦ صفات وابسته
۸۳ انواع مطالعه مداخله ای	۶۳	۱. همبستگی پیرسون (Correlation)
♦ ۱. RCT (Randomized clinical trial) یا	۶۳	۲. رگرسیون (Regression)
۸۳ کارآزمایی درمانی یا شاهددار تصادفی شده	۶۳	رگرسیون خطی
۸۳ ♦ ۲. Field trial یا کارآزمایی پروفیلاکتیک یا محلی	۶۳	سوالات آزمون‌ها
۸۳ ♦ ۳. کارآزمایی اجتماعی یا Community trials		
۸۴ مطالعات ثانویه یا Secondary review	۷۱	۷. برآورد، تعیین حجم نمونه و حدود اطمینان

۸۴	(Population Attributable Risk) PAR	۸۴	Narrative review .۱ ♦
	جز منتسب در گروه مواجهه (AF Attributable)	۸۴	Systemic review .۲ ♦
۸۵	(Fraction	۸۴	Review Article .۳ ♦
۸۶	(NNT) Number Needed to treat	۸۴	علیت (causality)
۸۶	مخدوش کننده‌ها	۸۴	نسبت‌های خطر
۸۷	روش‌های جلوگیری از مخدوش شدن		♦ اختلاف خطر (Risk Difference) یا خطرمنتسب
۸۷	سوالات آزمون‌ها	۸۴	(Attributable Risk)
			♦ خطر منتسب به جمعیت

فصل اول

متغیرها و نمودارهای توزیع فراوانی

متغیرها دو دسته هستند:

۱. کیفی

الف) اسمی (nominal): ضعیف ترین سطح اندازه گیری متغیرهاست. گروهها براساس نام و نشان مشخص شدند و نسبت به هم برتری و رتبه بندی ندارند. مانند جنسیت زن یا مرد، سفید یا سیاه، شهری یا روستایی

ب) رتبه ای یا ترتیبی (ordinal): گروهها براساس نام و نشان مشخص شدند و طبقات نسبت به هم برتری و رتبه بندی دارند ولی فاصله بین طبقات یکسان نیست. مانند: سطح تحصیلات (دیپلم، لیسانس، ... شدت موضوع (کم، متوسط، شدید)

۲. کمی

الف) فاصله ای (Interval): صفر قراردادی دارد. مقادیر منفی نیز میتواند داشته باشد. علاوه بر تقسیم بندی رتبه ای فاصله بین رتبهها نیز ثابت و مساوی است. اختلاف دو مقدار را میتوان به صورت فاصله بیان کرد. مانند: دما براساس سانتی گراد

ب) نسبتی (Ratio): صفر ذاتی دارد. مقادیر منفی را نیز شامل می شود. اختلاف دو مقدار را میتوان به

صورت نسبت بیان کرد. مقیاس نسبتی خود شامل دو گروه متغیرهای کمی پیوسته و کمی گسسته می باشد.

مقیاس کمی گسسته مقادیر مشخصی دارند که بین آنها مقدار دیگری نیست (اعشار نمی گیرد). مثل: تعداد نفرات.

مقیاس کمی پیوسته تمام اعداد حقیقی را میتواند شامل شود. مثل: وزن یا قد

در پایان یک مطالعه توصیفی بررسی فشارخون بالا در جمعیتی، فشارخون بالا به صورت فشار سیستولی بالای ۱۴۰ تعریف شد. متغیر فشار سیستولی در این بررسی از چه نوعی در نظر گرفته شده است؟

الف) کمی / نسبتی ب) کیفی / رتبه ای

ج) کمی / فاصله ای د) کیفی / اسمی

پاسخ: د

افراد به دو دسته تقسیم شدند:

۱) دارای BP بالای ۱۴۰

۲) دارای BP کمتر از ۱۴۰

نمودارها

۱. برای اطلاعات کیفی

الف) نردبانی (میله ای، Bar diagram) برای مقایسه اعداد در دوره های زمانی مختلف یا مکان های مختلف به کار می رود.

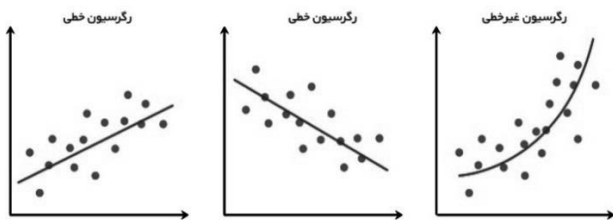
نمودار چند ضلعی یا پلی گون یا چندبر فراوانی، برای داده‌های کمی گسسته کاربرد دارد. اگر در نمودار هیستوگرام نقطه وسط هر مستطیل را به هم وصل کنیم نمودار پلی گون ایجاد می‌شود. (ج) خطی یا line graph: روند زمانی تغییرات را نشان می‌دهد.



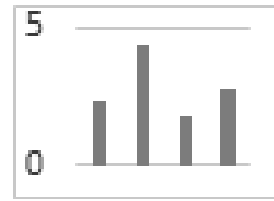
(د) شاخه برگگی یا ساقه و برگ Leaf and stem: برای نشان دادن همزمان توزیع و اصل مقادیر داده‌ها، اصل مقادیر در مقابل میزان فراوانی نوشته می‌شود.

2	5
3	
4	0
5	8
6	8 9 9
7	0 0 1 3 3 3 4 6 7 8
8	0 3 3 6 7
9	0 0 0 2 3 5 7
10	0 0

(ه) پراکنش (Dot plot): برای نشان دادن همبستگی دو متغیر کمی (ضریب همبستگی رگرسیون) به کار می‌رود.



(و) جدول توزیع فراوانی: در جدول توزیع فراوانی در یک ستون نام گروه و در ستون دیگر فراوانی مطلق، فراوانی نسبی و فراوانی تجمعی و... نوشته می‌شود.

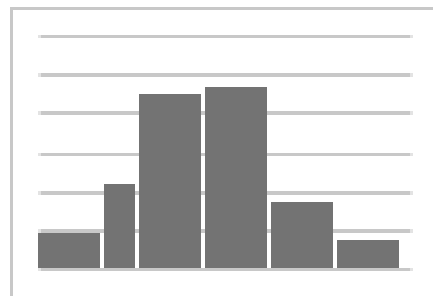


(ب) دایره ای (کیکی، Pia diagram)

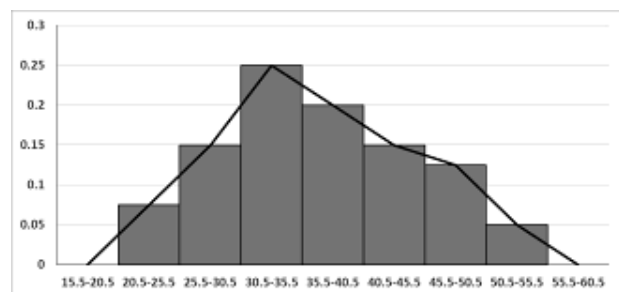
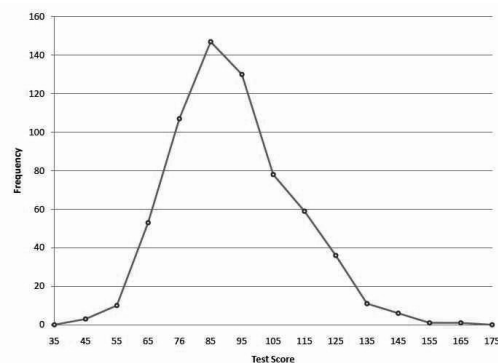


۲. برای اطلاعات کمی

(الف) هیستوگرام: یک نمودار میله ای است که برای توزیع فراوانی متغیرهای کمی پیوسته به کار می‌رود. هر مستطیل نشان دهنده فراوانی همان طبقه می‌باشد.



(ب) پلی گون = Poly gon = Scatter plot



فصل دوم

مرکزی ها

۱. میانگین یا Mean

عبارتست از جمع تمام داده‌ها تقسیم بر تعداد داده‌ها

$$\mu \text{ یا } m \text{ یا } \bar{x} = \frac{\sum x}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

میانگین از تمام داده‌ها تاثیر می‌گیرد.

۲. میانه یا Median

♦ شاخصی است که داده‌ها را به دو نیمه برابر تقسیم می‌کند در واقع نصف داده‌ها از آن بزرگتر و نصف دیگر از آن کوچکترند.

♦ اگر میانه را در نمودار هیستوگرام مشخص کنیم، سطح زیر هیستوگرام را به دو قسمت مساوی تقسیم می‌کند.

♦ برای داده‌های غیر متقارن و با انحراف زیاد، میانه همیشه وسط قرار گرفته و شاخص مناسب تری است.

♦ برای محاسبه میانه ابتدا اعداد را به ترتیب صعودی می‌نویسیم اگر تعداد داده‌ها فرد باشد، داده وسط معادل میانه است ولی در تعداد زوج میانگین دو داده وسطی معادل میانه است.

شاخص‌های مرکزی و شاخص‌های پراکندگی

شاخص‌های مرکزی

- ♦ میانگین
- ♦ میانه
- ♦ مد یا نما
- ♦ صدک و دهک

شاخص‌های پراکندگی

- ♦ دامنه (Range)
- ♦ دامنه میان چارکی (Interquartile Range)
- ♦ میانگین انحرافات (Mean deviation)
- ♦ انحراف معیار (Standard deviation)
- ♦ واریانس (Variance)
- ♦ ضریب تغییرات (CV)

$$\text{میانگین جدید} = 20 + 3 = 23$$

انحراف معیار بدون تغییر می ماند

$$\text{ضریب تغییرات} = \frac{\text{انحراف معیار}}{\text{میانگین}} = \frac{2}{23} = 8/7$$

۳- در بررسی تعداد فرزندان مادر شاغل در یک مرکز اطلاعات زیر به دست آمد. میانۀ تعداد فرزندان عبارت است از:

تعداد فرزندان	۰	۱	۲	۳	۴
فراوانی	۵	۸	۷	۴	۱

الف) ۱ ب) ۲ ج) ۸ د) ۷

پاسخ: الف

$$25 = 1 + 4 + 7 + 8 + 5 = \text{تعداد داده ها}$$

$$12.5 = 25 \div 2 = \text{محل میانه}$$

با توجه به اینکه تعداد کل نمونه ها برابر ۲۵ مورد است محل میانه داده وسطی یعنی ۱۲/۵ است که با توجه به جدول برابر یک می شود.

۴- در مطالعه ای فشار خون و کلسترول تعدادی از بیماران مبتلا به دیابت اندازه گیری شد. برای مقایسه پراکندگی فشار خون با کلسترول توتال این بیماران، کدام شاخص مناسب تر است؟ (پرانترنی اسفند ۹۲) الف) انحراف معیار ب) واریانس

ج) میانه و صدک

د) ضریب تغییرات

پاسخ: د

ضریب تغییرات برای داده های دارای واحدهای مختلف به کار می رود.

زمانی که تغییرپذیری در گروه های مشاهدات به میانگین اندازه گروه بستگی داشته باشد. برای مقایسه تغییرپذیری کدام شاخص مناسب است؟

الف) دامنه تغییرات ب) دامنه میان چارکی

ج) ضریب تغییرات د) میانه

پاسخ: ج

زمانی که تغییرپذیری گروه های مشاهدات به میانگین آن ها بستگی داشته باشد، ضریب تغییرات کاربرد دارد.

سوالات آزمون ها

۱- قرار است پراکندگی توزیع ضربان قلب با درجه حرارت بدن افراد یک جامعه با شاخصی مناسب مقایسه شود. کدام مورد زیر را توصیه می کنید؟ (پرانترنی اسفند ۹۰) الف) ضریب تغییرات

ب) انحراف معیار

ج) واریانس

د) دامنه

پاسخ: الف

ضریب تغییرات برای داده های دارای واحدهای مختلف به کار می رود یا به عبارتی برای مقایسه پراکندگی چند صفت یا یک صفت با چند واحد متفاوت به کار می رود.

۲- در انتهای یک مطالعه که به منظور برآورد میانگین وزن دانش آموزان انجام شده بود، مشخص گردید که ترازو وزن هر فرد را سه کیلوگرم کمتر از مقدار واقعی نشان داده است. اگر میانگین و انحراف معیار محاسبه شده به ترتیب ۲۰ و ۲ باشد مقدار ضریب تغییرات واقعی (پس از تصحیح وزن) چند درصد است؟ (پرانترنی شهریور ۹۱)

الف) ۱۰ ب) ۸/۷

ج) ۱۱/۵ د) قبل محاسبه نیست

پاسخ: ب

فصل سوم

احتمالات

قانون ضرب احتمالات

- هر وقت بین حوادث "و" داشته باشیم، قانون ضرب احتمالات استفاده می‌شود.
- احتمال وقوع همزمان چند پیشامد مستقل برابر حاصل ضرب تک تک هر کدام است.

$$P(A_1 A_2) = P(A_1) \times P(A_2)$$

- اگر پیشامدها مستقل از هم نباشند، در واقع پیشامد اول روی پیشامد دوم تاثیر داشته باشد:

$$P(A_1 A_2) = P(A_1) \times P(A_2 | A_1)$$

$P(A_2 | A_1)$: یعنی احتمال واقعه A_2 به شرطی که A_1 رخ داده باشد.

احتمال دختر بودن چهار فرزند متوالی می‌شود:

$$\frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

اگر از اعداد ۲ تا ۸ به صورت تصادفی انتخاب کنیم و جاگذاری نکنیم احتمال اینکه هر دو عدد زوج باشند برابر است با:

$$P(A_1 A_2) = P(A_1) \times P(A_2 | A_1) =$$

$$\frac{4}{7} \times \frac{3}{6} = \frac{2}{7}$$

$P(A_2 | A_1)$ یعنی A_2 به شرط A_1 : در واقع بعد از خروج عدد اول، یک عدد از تعداد کل (مخرج کسر) باید کم کرد.

قانون جمع احتمالات

- هر وقت حرف "یا" استفاده شود یا از کلمه "حداقل یکی" استفاده شود، قانون جمع احتمالات به کار می‌رود.

- احتمال جمع دو یا چند پیشامد ناسازگار برابر مجموع احتمالات تک تک آنهاست:

$$P(A_1 + A_2) = P(A_1) + P(A_2)$$

- حال اگر پیشامدها سازگار باشند (وقوع همزمان داشته باشند):

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1) \times P(A_2)$$

به دنبال یک عمل جراحی احتمال خونریزی ۵۰٪ و احتمال عفونت محل عمل ۲۰٪ است چقدر احتمال دارد فردی حداقل به یکی از این دو عارضه مبتلا شود؟

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1) \times P(A_2) =$$

$$\frac{50}{100} + \frac{20}{100} - \left(\frac{50}{100} \times \frac{20}{100}\right) = 60\%$$

تئوری بیز (Bayes)

تئوری توماس بیز روشی است برای یافتن احتمال دوم با استفاده از احتمال اول. تئوری بیز برای ارزیابی تست‌های تشخیصی و غربالگری به کار برده می‌شود. به کمک تئوری بیز می‌توان موارد زیر را محاسبه نمود:

حساسیت آزمون

♦ حساسیت نشان دهنده توانایی آزمون در شناسایی صحیح افراد بیمار می‌باشد. (حساسیت یعنی بیماریابی)

♦ حساسیت = احتمال تست مثبت در افراد بیمار

$$\text{حساسیت} = \frac{\text{بیماران با تست مثبت}}{\text{تعداد بیماران}}$$

♦ هرگاه شیوع یا کشندگی بیماری بالا باشد باید سطح حساسیت تست را افزایش داد.

ویژگی آزمون

♦ ویژگی تست = احتمال تست منفی در افراد سالم

$$\text{ویژگی} = \frac{\text{فرد سالم با تست منفی}}{\text{همه افراد سالم}}$$

کاربردها

♦ حساسیت برای R/O بیماری هاست (خاصیت screening دارد).

♦ ویژگی برای تایید تشخیص به کار می‌رود.

مثال: ANA، حساسیت بالایی برای تشخیص لوپوس دارد پس اگر منفی شود بیماری لوپوس رد می‌شود. اما زمانیکه مثبت شود برای تایید تشخیص یک تست دیگر با ویژگی بالا درخواست می‌کنیم.

۲۶- اگر احتمال اشتباه در تجویز هر دارو تنها ۵٪ باشد، در صورتیکه تعداد داروهای تجویزی به بیمار از ۳ دارو به ۵ دارو افزایش پیدا کند، احتمال تجویز اشتباه حداقل یک دارو چند درصد افزایش می‌یابد؟ (احتمال وقوع همزمان خطا را ناچیز بگیرد) (پرانتزنی اسفند ۹۷ - قطب ۱۰)

الف) ۲۰ (ب) ۱۰ (ج) ۱ (د) ۲
پاسخ: ب

هرگاه عبارت حداقل به کار رود از قانون جمع احتمالات استفاده می‌شود.

احتمال حداقل یک اشتباه برای سه دارو:

$$۰٫۵ + ۰٫۵ + ۰٫۵ = ۰٫۱۵ = ۱۵\%$$

احتمال حداقل یک اشتباه برای پنج دارو:

$$۰٫۵ + ۰٫۵ + ۰٫۵ + ۰٫۵ + ۰٫۵ = ۰٫۲۵ = ۲۵\%$$

$$۲۵\% - ۱۵\% = ۱۰\%$$

۲۷- دارویی با احتمال ۷۰ درصد یک بیماری را درمان می‌کند. چقدر احتمال دارد از بین ۲ بیماری که به طور تصادفی انتخاب شده‌اند و دارو را استفاده کرده اند، هیچ یک بهبود پیدا نکنند؟ (پرانتزنی آبان ۱۴۰۰)

الف) ۰٫۰۹ (ب) ۰٫۱۵ (ج) ۰٫۵۱ (د) ۰٫۴۹
پاسخ: الف

$$\text{احتمال عدم بهبودی: } ۰٫۰۹ = ۰٫۳ \times ۰٫۳$$

۲۸- دارویی با احتمال ۷۰ درصد یک بیماری را درمان می‌کند. چقدر احتمال دارد از بین ۲ بیماری که به طور تصادفی انتخاب شده‌اند و دارو را استفاده کرده اند، حداقل یک نفر بهبودی پیدا کند؟ (پرانتزنی اسفند ۱۴۰۰)

الف) ۰٫۰۹ (ب) ۰٫۹۱ (ج) ۰٫۵۱ (د) ۰٫۴۹
پاسخ: ب

$$\text{عدم بهبود هر دو مورد: } ۰٫۰۹ = ۰٫۳ \times ۰٫۳$$

$$\text{بهبود حداقل یک مورد: } ۱ - ۰٫۰۹ = ۰٫۹۱$$

۱۶- در یک جمعیت ده هزار نفری با شیوع ۱۰ درصد دیابت، آزمون غربالگری قند خون ناشتا صورت گرفت. اگر حساسیت آزمون ۷۰ درصد و ویژگی آن ۹۰ درصد باشد، تعداد بیماران برآورد شده چقدر است؟ (پرانتزنی اسفند ۹۳ - قطب همدان)

الف) ۹۰۰ نفر

ب) ۷۰۰ نفر

ج) ۱۶۰۰ نفر

د) ۱۰۰۰ نفر

پاسخ: ج

		Test	
		A	B
بیمار ☹️	+	C	D
	-		
		A+C=1000	D+B=9000
			سالم 😊

تعداد کل بیماران: $10000 \times 10\% = 1000$

$$\text{حساسیت} = \frac{\text{بیماران با تست مثبت}}{\text{تعداد بیماران}}$$

$$70\% = \frac{A}{1000} \rightarrow A = 700$$

تعداد کل افراد سالم: $10000 - 1000 = 9000$

$$D = 8100, B = 900, A + B = 1600$$

۱۷- در کدامیک از بیماری‌های زیر آزمون تشخیص با ویژگی (Specificity) بسیار بالا لازم است؟ (پرانتزنی شهریور ۹۴ - قطب آزاد)

الف) دیابت

ب) فشار خون

ج) ایدز

د) سل

پاسخ: ج

در بیماری‌هایی که تشخیص ندادن آنها عوارض جدی دارد مانند ایدز لازم است ویژگی تست را افزایش می‌دهیم.

۱۴- اگر در یک جمعیت ۲ هزار نفری، شیوع بیماری ۸٪ باشد و تست مورد استفاده برای غربالگری فقط در ۱۶ نفر از مبتلایان منفی شود و ۳۶۸ نفر از افراد سالم را به اشتباه دیابتی اعلام کند. حساسیت و ویژگی تست غربالگری چقدر است؟ (پرانتزنی شهریور ۹۳)

الف) حساسیت ۹۰٪، ویژگی ۸۰٪

ب) حساسیت ۸۰٪، ویژگی ۹۰٪

ج) حساسیت ۳۰٪، ویژگی ۷۰٪

د) حساسیت ۵۰٪، ویژگی ۲۰٪

پاسخ: الف

تعداد کل بیماران: $160 = 2000 \times 8\%$

تعداد بیمار تست مثبت: $160 - 16 = 144$

$$\text{حساسیت} = \frac{\text{بیمار با تست مثبت}}{\text{تعداد بیماران}} = \frac{144}{160} = 90\%$$

تعداد کل افراد سالم: $1840 = 2000 - 160$

افراد سالم با تست منفی: $1472 = 1840 - 368$

$$\text{ویژگی} = \frac{\text{افراد سالم با تست منفی}}{\text{افراد سالم}} = \frac{1472}{1840} = 80\%$$

۱۵- در صورتیکه ۶۰ درصد زنان دارای ماموگرافی مثبت، مبتلا به سرطان پستان باشند. این شاخص غربالگری چه نامیده می‌شود؟ (پرانتزنی اسفند ۹۳ - قطب ۷)

الف) حساسیت

ب) ارزش اخباری مثبت

ج) میزان مثبت کاذب

د) ویژگی

پاسخ: ب

ارزش اخباری مثبت = احتمال بیمار بودن در یک فرد با تست مثبت

فصل چهارم

توزیع احتمالات

توزیع پواسون

مواردی وجود دارد که تعداد رخداد یک پدیده قابل شمارش می‌باشد اما تعداد عدم رخداد را نمی‌توان شمرد، در این موارد توزیع پواسون به کار می‌رود. مثال: تعداد تصادفات در ماه قابل شمارش است ولی تعداد موارد عدم وقوع تصادف خیر.

- برای داده‌های کمی گسسته کاربرد دارد.
- در توزیع دو جمله ای که n در آن بسیار بزرگ و p در آن بسیار کوچک است (مثل بیماری نادری که احتمال وقوع آن بسیار کم است)
- میانگین و واریانس در توزیع پواسون برابر هستند

$$\text{فرمول پواسون} = \frac{e^{-m} \times m^x}{x!}$$

m : میانگین، x : تعداد رخداد

شیوع موارد یک بیماری ژنتیکی یک در هزار است (۰/۰۰۱) احتمال آنکه در نمونه ۱۰۰۰۰ نفری، هیچ موردی از این بیماری پیدا نشود را با کدام توزیع بهتر میتوان محاسبه کرد؟

الف) توزیع نرمال (ب) دو نمایی (Bi modal)
ج) پواسون (د) دو جمله ای
پاسخ: ج

احتمال تصادف در یک بزرگراه دارای توزیع پواسون با میانگین ۳ تصادف در ماه است. احتمال اینکه در یک ماه تصادفی رخ ندهد چقدر است؟

$$m=3, x=0$$

$$\frac{e^{-m} \times m^x}{x!} = \frac{e^{-3} \times 1}{0!} = e^{-3}$$

سوالات آزمون ها

۱- در توزیع دو جمله ای اگر حجم نمونه بزرگ (a) و نسبت موفقیت کوچک (p) باشد طوری که میانگین آن (ap) کمتر از ۵ باشد آنگاه به جای توزیع دو جمله ای از کدام توزیع می‌توان تقریب زد؟ (پرانترنی اسفند ۹۳- قطب ۲)

الف) نرمال

ب) هندسی

ج) فول هندسی

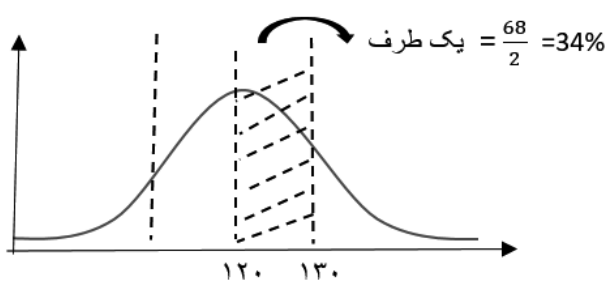
د) پواسون

پاسخ: د

۲- احتمال تصادف در یک بزرگراه دارای توزیع پواسون با میانگین ۳ تصادف در ماه است. احتمال اینکه در یک ماه تصادفی رخ ندهد چقدر است؟ (پرانترنی شهریور ۹۳ - قطب ۱۰)

الف) e^{-3} (ب) $1 - e^{-3}$ (ج) $3e^{-3}$ (د) $1 - 3e^{-3}$

پاسخ: الف



۲- میانگین قندخون برابر ۱۰۰ و واریانس آن ۱۶ می‌باشد. محدوده نرمال قند خون جامعه چقدر است؟

واریانس = ۱۶

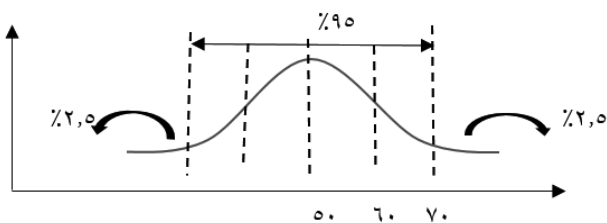
انحراف معیار = $\sqrt{16} = 4$

اختلاف دو انحراف معیار ۲ SD محدوده نرمال پزشکی محسوب می‌شود.



پس محدوده نرمال ۹۲ تا ۱۰۸ می‌شود.

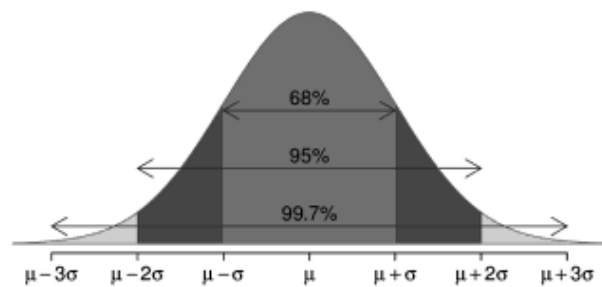
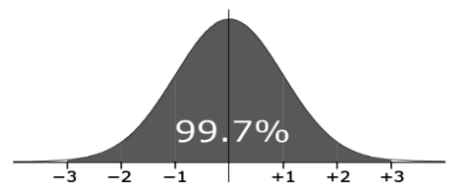
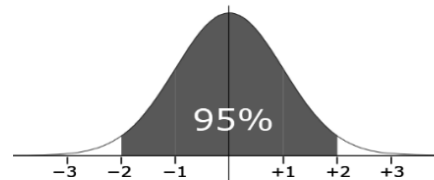
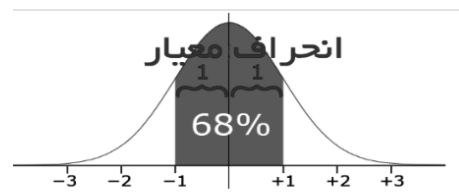
۳- در آزمونی که ۱۰۰۰ نفر شرکت کننده داشته است، میانگین نمرات شرکت کنندگان ۵۰ و انحراف معیار آن ۱۰ به دست آمده است. در صورتیکه توزیع نرمال داشته باشد. چند نفر از شرکت کنندگان نمره بالای ۷۰ دارند؟



افراد دارای نمره بالای ۷۰ معادل ۲,۵% جامعه هستند

پس تعداد آنها برابر است با: $2,5\% \times 1000 = 25$

اعداد و درصدها مهم هستند و باید به خاطر بسپارید.



- ♦ توجه کنید که به طور مثال در فاصله یک انحراف معیار از میانگین ۳۴% داده‌ها در یک طرف و ۳۴% آنها در طرف دیگر قرار دارند. ($68 \div 2 = 34$)
- ♦ بازه ۹۵% یعنی اختلاف دو انحراف معیار محدوده نرمال پزشکی محسوب می‌شود.

نمونه سوالات تمرینی

۱- میانگین فشار خون ۱۲۰ با انحراف معیار ۱۰ می‌باشد چند درصد جامعه درای فشارخون بین ۱۲۰ تا ۱۳۰ هستند؟

۱۵- در یک توزیع نرمال مساحت بین مقدار میانگین به علاوه و منهای ۲ انحراف معیار برابر است با: (پرانترنی شهریور ۹۴ - قطب آزاد)

الف) ۳۴ درصد

ب) ۶۸ درصد

ج) ۹۵ درصد

د) ۹۸ درصد

پاسخ: ج

۱۶- اگر یک شاخص خونی در نوزادان دارای توزیع نرمال با میانگین ۱۲ و انحراف معیار ۳ باشد، تقریباً چند درصد از نوزادان شاخص خونی بیشتر از ۶ دارند؟ (پرانترنی اسفند ۹۴ - قطب ۷)

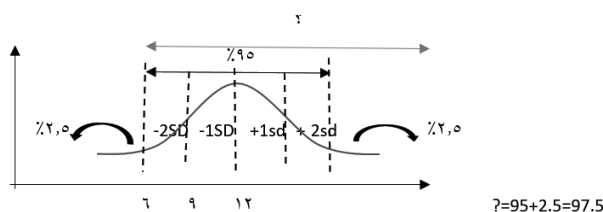
الف) ۹۷٫۵

ب) ۹۵

ج) ۵

د) ۲٫۵

پاسخ: الف



۱۷- فرض کنید متغیر X در یک جامعه دارای توزیع نرمال می‌باشد. اگر میانگین و واریانس X به ترتیب ۱۵۰ و ۲۲۵ باشد. تقریباً ۶۸ درصد افراد در چه محدوده ای قرار می‌گیرند؟ (پرانترنی اسفند ۹۴ - قطب ۹)

الف) ۱۲۰ - ۱۸۰

ب) ۱۳۵ - ۱۶۵

ج) ۱۵۰ - ۱۶۵

د) ۱۰۵ - ۱۹۵

پاسخ: ب

$15 = \sqrt{225} = \text{واریانس} = \text{انحراف معیار}$

$135 \text{ تا } 165 = 150 \pm 15 = 150 \pm 1sd$

۱۲- اگر توزیع داده‌ها متقارن نباشد رابطه شاخص‌های مرکزی به صورت زیر است؟ (پره شهریور ۹۴ - قطب ۳)

الف) میانگین بین میانه و Mode قرار می‌گیرد.

ب) میانه بین میانگین و Mode قرار می‌گیرد.

ج) Mode بین میانگین و میانه قرار می‌گیرد.

د) تاثیری بر مقدار این سه شاخص ندارد.

پاسخ: ب

در توزیع غیرمتقارن (دارای چولگی) میانه همواره بین میانگین و مد قرار می‌گیرد.

۱۳- اگر توزیع فراوانی قد نوزادان نرمال با میانگین ۵۰ سانتی‌متر و واریانس ۹ باشد، آنگاه تقریباً چند درصد از نوزادان این جامعه دارای قد بالاتر از ۵۶ سانتی‌متر هستند؟ (پرانترنی شهریور ۹۴ - قطب ۷)

الف) ۳۴

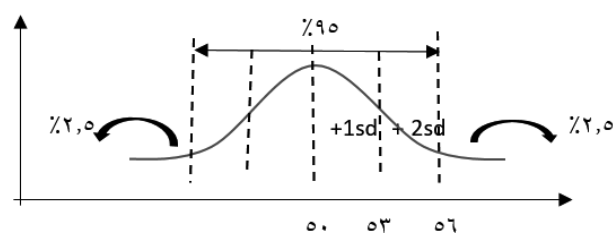
ب) ۵

ج) ۲٫۵

د) ۰٫۰۵

پاسخ: ج

$3 = \sqrt{9} = \text{واریانس} = \text{انحراف معیار}$



۱۴- در یک توزیع نرمال استاندارد همواره واریانس صفت مورد بررسی: (پرانترنی شهریور ۹۴ - قطب ۹)

الف) بزرگ‌تر از میانگین است.

ب) کوچکتر از میانگین است.

ج) برابر جذر میانگین است.

د) برابر میانگین است.

پاسخ: الف

فصل پنجم

اصول تفسیر آماری

آزمون فرضیه

فرضیه صفر (H_0)

هر اختلاف قابل مشاهده ای تنها در اثر خطای نمونه برداری ایجاد شده و تصادفی است در واقع تفاوت حقیقی در جمعیت مورد مطالعه نداریم.

- بین محاسبه شما و فرضیه تفاوت و رابطه ای نیست.

H_0 = صفر بی تفاوت (رابطه ندارد)

فرضیه H_1

- یعنی تفاوت معنادار یا ارتباط وجود دارد. (برعکس قبلی)

انواع خطاها

- در هنگام آزمون فرضیه با دو نوع خطای آماری مواجه می شویم:
- خطای نوع اول یا آلفا (α): یعنی رد H_0 در حالی که درست است.
- خطای نوع دوم یا بتا (β): یعنی اختلاف هست ولی تشخیص داده نمی شود. (برعکس قبلی است)

نکته: آلفا را بخوانید رد درست (یعنی تفاوت نیست ولی اشتباها می گه هست) بتا می شود: تایید به غلط

H_0 (عدم اختلاف واقعیت)
غلط ← درست ←

پذیرش تصمیم	رابطه ای نیست (تفاوتی نیست)	خطای β
رد تصمیم	خطای α	Power

نکته: نمی توان هر دو خطا را همزمان کاهش داد چون اگر یکی کاهش یابد دیگری افزایش پیدا می کند.

اکثراً "خطای نوع اول در یک حد، ثابت می شود و خطای نوع دوم به حداقل رسانده می شود. بعد از ثابت شدن خطای نوع اول می توان خطای β را با افزایش حجم نمونه کم کرد.

توان آزمون یا Power

$1 - \beta$ = توان آزمون

- یعنی احتمال اینکه فرضیه صفر رد شود در حالی که غلط است.
- یعنی H_0 غلط بوده و ما قدرت داشتیم آن را رد کنیم.

سوالات آزمون ها

۴- کدام عبارت درست نیست؟ (پرانترنی شهریور ۹۴ - قطب ۳)

- (الف) فرض صفر ممکن است در سطح ۱ درصد رد شود ولی سطح ۵ درصد رد نشود.
 (ب) در صورتی که فرض صفر درست باشد و آن را رد کنیم مرتکب خطای نوع اول شده ایم.
 (ج) با ثابت ماندن سایر شرایط، با افزایش حجم نمونه احتمال رد فرض صفر افزایش می‌یابد.
 (د) فرض صفر ممکن است در سطح ۵ درصد رد شود ولی در سطح یک درصد رد نشود.
 پاسخ: ب

۵- اگر در آزمون یک فرضیه، سطح اشتباه نوع اول $(\alpha = 0.05)$ انتخاب شود کدامیک از موارد زیر صحیح است؟ (پرانترنی شهریور ۹۴ - قطب ۴)

(الف) احتمال قبول H_0 وقتی که H_0 واقعا غلط است ۱٪ است.
 (ب) احتمال رد H_0 وقتی که H_0 واقعا درست است ۱٪ است.
 (ج) احتمال اینکه نتیجه آزمون فرض معنی دار شود ۱٪ است.
 (د) فرض H_0 حداقل در ۱٪ موارد صحیح است.
 پاسخ: ب
 خطای آلفا یعنی رد درست (رد فرضیه درحالیکه درست است)

۶- در آزمون فرضیه $\mu_1 = \mu_0$: H_0 در مقابل $\mu_1 < \mu_0$ ، مقدار α (احتمال خطای نوع اول) بیان کننده میزان احتمال: (پرانترنی اسفند ۹۴ - قطب ۲)

(الف) نادرست بودن فرضیه صفر است.
 (ب) نادرست بودن فرضیه مقابل است.
 (ج) رد فرض مقابل وقتی که این فرضیه صحیح است.
 (د) رد فرض صفر، وقتی این فرضیه صحیح است.
 پاسخ: د

۱- پژوهشگری جهت محاسبه توان آزمون خود نیاز به دانستن کدامیک از مقادیر زیر دارد؟ (پرانترنی شهریور ۹۳ - قطب ۳)

(الف) مقدار d.f یا درجه آزادی
 (ب) مقدار n یا حجم نمونه
 (ج) سطوح اطمینان
 (د) مقدار β یا خطای نوع دوم
 پاسخ: د

$\beta = 1 - \text{توان آزمون}$

۲- مفهوم خطای نوع دوم (بتا) با کدام گزینه همخوانی دارد؟ (پرانترنی شهریور ۹۳ - قطب ۴)

(الف) پذیرش H_1 با وجود صحت H_0
 (ب) پذیرش H_0 با وجود صحت H_1
 (ج) رد H_0 با وجود صحت H_0
 (د) رد H_1 با وجود صحت H_1
 پاسخ: الف

۳- در یک آزمون فرضیه وقتی فرضیه H_0 در سطح معنی داری α رد شود که مقدار P (P - Value) محاسبه شده: (پرانترنی شهریور ۹۴ - قطب ۲)

(الف) کمتر از α باشد.
 (ب) بزرگتر از $1 - \alpha$ باشد.
 (ج) بزرگتر از α باشد.
 (د) کمتر از $1 - \alpha$ باشد.
 پاسخ: الف
 اگر P - value کمتر از α باشد = تفاوت معنی دار است (H_0 رد می‌شود)

فصل ششم

آزمون کای اسکوئر و آزمون‌های معنادار

برای انتخاب نوع آزمون آماری مناسب براساس نوع متغیرها و داده‌ها عمل می‌شود.

۱. اطلاعات کمی

الف) مقایسه میانگین دو گروه با هم

اگر $n < 30$ ← Z test

اگر $n > 30$ ← T test

$$T = \frac{|\mu_1 - \mu_2|}{SEM} = \frac{|\mu_1 - \mu_2|}{\frac{\delta}{\sqrt{n}}}$$

نکته: وقتی حجم نمونه زیاد شود، توزیع t، شبیه توزیع نرمال می‌شود.

ب) مقایسه میانگین بیش از دو گروه

ANOVA = آنالیز واریانس = F test

ج) مقایسه میانگین دو جز وابسته

Pair T - test (آزمون t زوجی)

Pair Z - test

مانند قبل و بعد از یک اتفاق eg قبل و بعد رژیم

غذایی یا تفاوت بین دو گروه زوجی eg میانگین قد دوقلوها

۲. اطلاعات کیفی

♦ برای مقایسه نسبت‌ها، درصدها یا شیوع در گروه‌های مختلف، مناسب هستند

الف) مستقل

کای اسکوئر $\chi^2 = \text{chi squar}$

♦ برای مقایسه تاثیر دو درمان مختلف یا برای نسبت مشاهدات در گروه‌های مختلف به کار می‌رود.

= درجه آزادی کای اسکوئر

$$(r - 1) \times (c - 1) = (\text{تعداد ستون} - 1) \times (\text{تعداد ردیف} - 1)$$

ب) غیر مستقل Mc Nemor

نکات آزمون کای اسکوئر (chi squar) یا کای دو χ^2 :

♦ برای بستگی اطلاعات کیفی اسمی (Nominal) به کار می‌رود.

♦ آزمون کای دو براساس تعداد واقعی موارد مشاهده شده در هر گروه است نه براساس نسبت یا درصد.

♦ وقتی می‌خواهیم ۲ یا چند نسبت مستقل را در ۲ یا چند گروه مقایسه کنیم، کاربرد دارد.

♦ آزمون نسبت‌ها می‌باشد در واقع می‌گوید آیا نسبت مشاهدات که در گروه‌های مختلف قرار

یک پژوهشگر با انجام مصاحبه تلفنی با نمونه ۹۰۰ نفره تصادفی از دانشجویان، تست استعمال دخانیات را ۱۰٪ برآورد کرده است دامنه اطمینان ۹۵٪ برای نسبت مذکور چقدر است؟

الف) ۴ - ۱۶

ب) ۷ - ۱۳

ج) ۸ - ۱۲

د) ۹ - ۱۱

شیوع $p = 10\%$

$$n = \frac{Z^2 p (1-p)}{d^2}$$

$$d = z \sqrt{\frac{p(1-p)}{n}} = 2 \times \sqrt{\frac{0.1 \times 0.9}{900}} = 2 \times 0.01 = 0.02 = 2\%$$

$$CI = \bar{x} \pm d = 10\% \pm 2\% = 8\% \text{ تا } 12\%$$

در برآورد میانگین یک صفت اگر حجم نمونه را ۲ برابر کنیم چه تاثیری در فاصله اطمینان رخ می‌دهد؟

الف) دوبرابر می‌شود

ب) نصف می‌شود

ج) ۴ برابر می‌شود

د) ۲ برابر کاهش می‌یابد

$$CI = \bar{x} \pm d \quad d = \frac{z \delta}{\sqrt{n}}$$

اگر حجم نمونه دو برابر شود:

$$CI = \bar{x} \pm \frac{1}{\sqrt{2}} d$$

پاسخ: د

افزایش کدامیک از موارد زیر منجر به نزدیک‌تر شدن حد بالا و حد پایین برآورد فاصله ای میانگین، به همدیگر می‌شود؟

الف) میانگین (ب) حجم نمونه

ج) انحراف معیار (د) درصد اطمینان

پاسخ: ب

شیوع یک بیماری حدود ۱۰٪ می‌باشد. حجم نمونه ای با ۹۵٪ اطمینان این شیوع را با خطای کمتر از ۲٪ برآورد می‌کند چقدر است؟

$$n = \frac{Z^2 p (1-p)}{d^2} = \frac{2^2 \times 0.1 \times 0.9}{(0.02)^2} = 900$$

حدود اطمینان میانگین جامعه

• یعنی با اطمینان ۹۵٪ میتوانیم بگوییم میانگین جمعیت در این محدوده است. (با فرض α درصد ۵ درصد)

$$CI = \bar{x} \pm d$$

\bar{x} : میانگین نمونه‌ها

d: دقت برآورد (اشتباه برآورد) (معمولا از فرمول حجم به دست می‌آید)

α : خطا

در یک نمونه از کودکان ۵ ساله با میانگین قد ۱۰۰ سانتی‌متر و خطای استاندارد ۰٫۵ حدود اطمینان ۹۵٪ چقدر است؟

$$SEM = \frac{\delta}{\sqrt{n}} = 0.5$$

$$n = \frac{Z^2 \delta^2}{d^2} \rightarrow d = \frac{Z \delta}{\sqrt{n}}$$

$$z \times \frac{\delta}{\sqrt{n}} = z \times sem = 2 \times 0.5 = 1$$

حدود اطمینان برای شیوع (نسبت)

$$CI = P \pm d$$

P: شیوع، d: خطای برآورد

فصل هشتم

انواع مطالعات

(۱) توصیفی

مانند Case report / Case series فقط گزارش موارد است و هیچ مقایسه ای انجام نمی‌شود.

(۲) تحلیلی یا آنالیتیک

مداخله دارد (بالینی‌ها)

مانند: RCT (Random clinical trial)

مداخله ندارد:

۱- محیطی یا تطبیقی یا اکولوژیک

۲- مقطعی یا مطالعه شیوع یا Cross sectional

۳- مورد شاهدهی یا Case control

۴- کوهورت یا همگروهی

در مطالعات توصیفی در انتها به فرضیه می‌رسیم درحالی‌که در مطالعات تحلیلی، شروع مطالعه براساس فرضیه است.

مطالعات مقایسه ای براساس دو عامل دسته بندی می‌شوند:

♦ مواجهه (exposure) یا risk factor

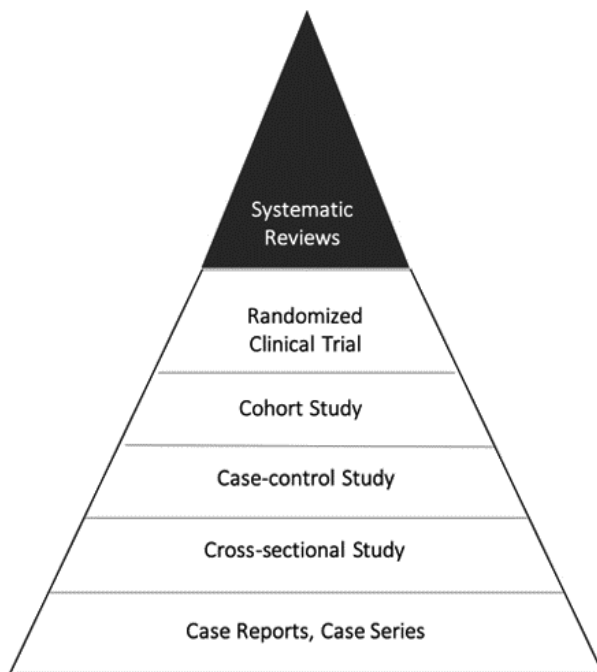
♦ بیماری یا out come

♦ مطالعه مقطعی یا شیوع (Cross sectional):

در شروع مطالعه همزمان RF و Out come را

بررسی می‌کنیم.

- ♦ مطالعه مورد شاهدهی (Case control): از بیماری یا out come به سمت مواجهه یا RF می‌رویم.
- ♦ کوهورت و بالینی (Clinical trial): از RF یا مواجهه به بیماری می‌رسیم.
- ♦ در مطالعه Case control با Odd ratio (OR) به رابطه بین پیامد و علت پی می‌بریم.
- ♦ در مطالعه کوهورت با Relative Risk (RR) به رابطه علت معلولی پی می‌بریم.
- ♦ در این تصویر ارزش مطالعات به صورت هرم نمایش داده شده است. هرچه به سمت بالای هرم حرکت کنیم ارزش مطالعه بیشتر شده و تعیین رابطه علت معلولی قوی تر می‌شود:



Source: Adapted from Ho et al. 2008¹⁰.

مطالعات ثانویه یا Secondary review

۱. Narrative review

نتیجه مطالعات مختلف را کنار هم شرح می‌دهد ولی جمع بندی و نتیجه گیری نهایی انجام نمی‌دهد.

۲. Systemic review

نتیجه مطالعات را شرح می‌دهد آنالیز آماری هم انجام می‌دهد و در نهایت به یک نتیجه می‌رسد.
 ♦ برای انجام این مطالعه نیاز به داشتن تخصص در آن رشته نیست و توسط کارشناس آمار قابل انجام است در نتیجه خیلی دقیق نیست.
 ♦ روش آماری که استفاده می‌شود، متاآنالیز است.

۳. Review Article

نتایج مطالعات را کنار هم شرح می‌دهد و در نهایت نظر خود را اعلام می‌کند.
 ♦ فقط توسط استاد برجسته که در آن موضوع تخصص دارد قابل انجام است.

علیت (causality)

برای اثبات وجود رابطه علیتی باید معیارهای Bradford Hill برقرار باشد:
 ۱- تقدم زمانی (Temporality): مهمترین معیار تقدم عامل خطر بر بروز بیماری است.
 ۲- قدرت ارتباط (Strength of association): هرچه RR بالاتر قدرت ارتباط بیشتر است.
 ۳- ثبات یا پایداری ارتباط یا تکرارپذیری (Consistency): با افراد مختلف در مکان‌ها و نمونه‌های مختلف رابطه را اثبات کرده است.
 ۴- توجیه پذیری یا منطقی بودن از نظر علمی (Plausibility)
 ۵- رابطه دوز - پاسخ Dose - Response: با افزایش دوز، RR بالاتر برود.
 ۶- اختصاصی بودن (Specificity): بیماری فقط در

اثر یک عامل ایجاد شود.

۷- پیوستگی ارتباط یا همانندی (Analogy): اطلاعات قبلی با اطلاعات به دست آمده مطابقت داشته باشد.

۸- طراحی مطالعه یا مداخله (Study design or Experiment)

در صورتیکه شواهد مطالعه تجربی شاهددار برای تعیین علیت وجود نداشته باشد همه موارد زیر می‌توانند، برای رد بررسی احتمال وجود رابطه علیتی برای ارتباط مشاهده شده موثر باشند به جز:

- الف) شدت ارتباط
- ب) معنی داری آماری
- ج) خطر نسبی
- د) تقدم زمانی
- پاسخ: ب

نسبت‌های خطر

اختلاف خطر (Risk Difference) یا خطرمنتسب (Attributable Risk)

یعنی تفاوت میزان وقوع بین گروه دارای مواجهه و گروه بدون مواجهه.
 لذا خطر اضافه ناشی از مواجهه با عامل خاص را مشخص می‌نماید.
 در موارد بیماری‌های حاد و در ابتدای شروع گسترش بیماری کاربرد دارد.

بروز در مواجهه نیافته - بروز در مواجهه یافته = AR

$$AR = \frac{A}{A+B} - \frac{C}{C+D}$$

خطر منتسب به جمعیت

(Population Attributable Risk) PAR

بیان می‌کند خطر اضافی بیماری در کل جمعیت

۵. استاندارد کردن میزان‌ها (Standardization of rates)

۶. روش‌های Multivariate: برای اصلاح متغیرهای وابسته به یک یا چند عامل مستقل مخدوش گر می‌توان از رگرسیون مثل رگرسیون لجستیک، رگرسیون COX یا آنالیز کواریانس استفاده نمود.

۷. روش‌های Post Analytic: در مرحله تجزیه و تحلیل اطلاعات به کار می‌رود. شامل لایه بندی (Stratification) و الگوبرداری آماری (Statistical Modelling) می‌شود.

سوالات آزمون‌ها

۱- همه موارد زیر از مزایای انجام مطالعه کوهورت می‌باشند به جز: (پرانتزنی ۸۶)

الف) می‌توان بروز را محاسبه نمود.

ب) نسبت‌های مقدار پاسخ قابل محاسبه هستند.

ج) جهت بررسی عوامل خطر بیماری نادر مناسب هستند.

د) می‌توان برآورد مستقیمی از خطر نسبی ارائه دهد.

پاسخ: ج

۲- درنظراست مطالعه ای به منظور بررسی ارتباط بین انسفالوپاتی اسفنجی گاوی (BSE) با مصرف نوع خاصی از گوشت فرآوری شده، سنجیده شود. با توجه به ویژگی‌های بیماری مزبور طراحی مطالعه به چه شکلی منطقی تر است؟ (پیش‌کارورزی - شهریور ۹۰)

الف) مورد - شاهدهی

ب) همگروهی تحلیلی

ج) مقطعی تحلیلی

د) کارآزمایی بالینی

پاسخ: الف

مطالعه مورد-شاهدهی، برای بررسی علل بیماری‌های نادر، بیماری‌هایی که بروز پایین دارند بهترین روش است.

♦ خطای اندازه‌گیری (Measurement): شامل:

۱- خطای یادآوری (Recall) ۲ - سوگرایی مشاهده‌گر (Observer Bias) که راه برطرف کردن با Blind کردن است.

خطای Confounding: وجود هم زمان متغیرهایی که هم با عامل خطر و هم با پیامد (بیماری) مرتبط هستند. هم زمانی دو یا چند عامل با هم، مثل مصرف سیگار و الکل به صورت هم زمان.

♦ تورش اطلاعات (information): افراد از مواجهه با عامل خطر مطلع باشند.

نکته: دو تا بایاس مربوط به مطالعه Case Control می‌شود. خطای انتخاب و خطای یادآوری خطای انتخاب: اختلاف منظم بین ویژگی افراد انتخاب شده و انتخاب نشده. مثل انتخاب افرادی که به درمانگاه خاصی مراجعه می‌کنند یا خطای کارگر سالم.

خطای یادآوری: موردها، مواجهه را بهتر به یاد می‌آورند یا بیشتر و کمتر تخمین می‌زنند.

روش‌های جلوگیری از مخدوش شدن

۱. تصادفی سازی (Randomization): فقط در مطالعات مداخله ای (بالینی و تجربی) کاربرد دارد.

۲. همسان سازی (Matching): بیشتر در مطالعه Case Control به کار می‌رود. به طور مثال سن و جنس دو گروه یکسان باشد.

۳. محدود سازی: (Restriction) محدودسازی به افرادی که خصوصیت ویژه ای دارند (عاملی که شک دارند را کنار گذارند) مثل انتخاب سیگاری‌های غیر الکلی

۴. طبقه بندی (Stratification): گروه بندی افراد تحت مطالعه

نکته: موارد ۲ تا ۴ در مطالعات مشاهده ای کاربرد دارند.

$$RR = \frac{\text{وقوع در افراد با RF مثبت} + \text{بروز در مواجه یافتگان}}{\text{وقوع در افرادی که در معرض RF نبودن} + \text{بروز در مواجه نیافتگان}} = \frac{100}{\frac{1000}{10}} = 5$$

۶۸- تورش یادآوری (Recall Bias) در کدامیک از مطالعات اپیدمیولوژیک باعث سوگرایی بیشتری می‌گردد؟ (دستیاری ۹۴)
 الف) Case - Control
 ب) Historical Cohort
 ج) Cross Sectional
 د) Prosoective Cohort
 پاسخ: الف

۶۹- در یک مطالعه موردشاهدی که به بررسی ارتباط سیگار با سرطان مثانه می‌پردازد. نتایج به صورت جدول زیر به دست آمده است. شاخص مناسب برای تعیین ارتباط بین مواجهه و بیماری کدام گزینه صحیح است؟ (پرانترنی اسفند ۹۴ - قطب ۳)

	سرطان مثانه	گروه شاهد
مصرف سیگار	۱۴۰	۴۰
عدم مصرف سیگار	۶۰	۱۶۰

الف) خطر نسبی = ۲/۹
 ب) نسبت شانس = ۲/۹
 ج) خطر نسبی = ۹/۳
 د) نسبت شانس = ۹/۳
 پاسخ: د
 در مطالعه مورد شاهدی، نسبت شانس یا odd ratio محاسبه می‌شود.

$$\text{Odd Ratio (OR)} = \frac{\text{شانس بیماری در افراد مواجه یافته}}{\text{شانس بیماری در افراد مواجه نیافته}} = \frac{A \times D}{B \times C} = \frac{A(++) \times D(--)}{B \times C} = \frac{140 \times 160}{60 \times 40} = 9,3$$

۶۵- منظور از تخصیص تصادفی (Random allocation) در کارآزمایی‌های بالینی کدام است؟ (پرانترنی اسفند ۹۴ - قطب ۶)
 الف) افراد به دلخواه خود در گروه‌های مداخله و کنترل قرار گیرند.
 ب) شانس ورود افراد به گروه‌های مداخله و کنترل یکسان باشد.
 ج) افراد، تعلق خود را به گروه‌های مداخله و کنترل ندانند.
 د) افراد گروه‌های مداخله و کنترل با توافق محقق در گروه‌ها قرار گیرند
 پاسخ: ب

۶۶- در یک مطالعه موردشاهدی، شانس مواجهه با عامل خطر احتمالی، در افراد مبتلا به بیماری چهار برابر افراد سالم بوده است. کدام فاصله اطمینان (CI) ارائه شده دلالت بر معناداری این رابطه ندارد؟ (دستیاری ۹۵)
 الف) ۶ - ۲
 ب) ۶ - ۱/۵
 ج) ۶ - ۰/۵
 د) ۰/۲۵ - ۰/۵
 پاسخ: ج

حدود اطمینان RR و OR نباید شامل یک باشد و اگر شامل یک باشد یعنی آن حدود اطمینان معنی دار نیست.

۶۷- نتایج ۱۰ ساله پیگیری یک مطالعه همگروهی دو گروه مواجهه یافته با آزیست و مواجهه نیافته با آزیست در بروز سرطان ریه به شرح زیر می‌باشد. خطر نسبی مواجهه با آزیست در بروز سرطان ریه کدام است؟ (پرانترنی اسفند ۹۴ - قطب ۱)

مواجهه با آزیست	سرطان		جمع
	دارد	ندارد	
دارد	۱۰۰	۹۰۰	۱۰۰۰
ندارد	۱۰	۴۹۰	۵۰۰

الف) ۲۰
 ب) ۱۰
 ج) ۵
 د) ۴
 پاسخ: ج