

Course on :

Data Analysis in

CASE CONTROL Studies

دوره آموزشی:

تحلیل آماری در

مطالعات مورد شاهدهی



Exposed



Not exposed

آغاز دوره از
سه شنبه ۲۹ مهر

۱۸ ساعت
آموزش
نظری و
عملی

طراحی و تدریس: دکتر عباس کشتکار

عضو هیات علمی دانشگاه علوم پزشکی تهران

مدرس بیش از ۱۵۰ عنوان دوره و کارگاه آموزشی برای پژوهشگران

زمان جلسات

سه شنبه شبها: ۸ تا ۱۱

گامهای تحلیل آماری در مطالعات مورد شاهدهی :

1. تحلیلهای توصیفی یا Descriptive Analysis

2. تحلیل تک متغیره یا Univariable A.

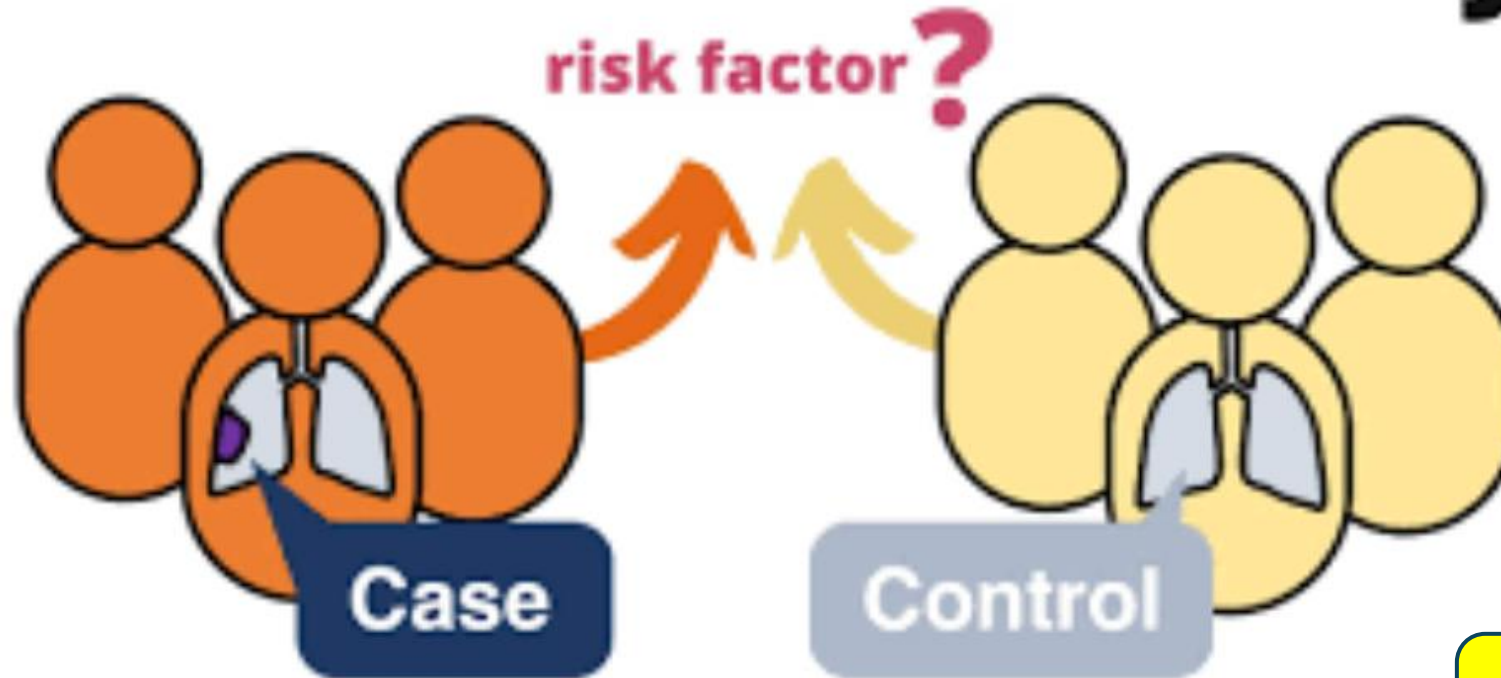
3. تحلیل چند متغیره یا Multivariable A.

4. تحلیلهای اضافی یا Additional A شامل:

(a) تحلیلهای زیر گروه یا Sub-group A.

(b) تحلیلهای حساسیت یا Sensitivity A.

Case-control study



درس سوم

تحليل توصيفي و تک متغيره در
مطالعات مورد شاهدهی

آیا تحلیل توصیفی در مطالعات مورد شاهدهی، ارائه میگردد؟

- با توجه به ماهیت «غیرتوصیفی» یا «تحلیلی» مطالعات مورد شاهدهی، پاسخ خیر است.
- فلذا از بکاربردن شاخصهای توصیفی مانند شیوع و ... در مطالعات مورد شاهدهی، چه بصورت کلی و چه بصورت تفکیک در گروههای مورد و شاهد، پرهیزید.
- در واقع، حتی اگر از شاخصهای ساده توصیفی از قبیل «میانگین» یا «فراوانی»، بصورت مقایسه‌ای استفاده میشود (بصورت کل افراد تحت مطالعه یا ترکیب مورد و شاهد، کاربرد و استفاده‌ای ندارد)
- یکی از بهترین ابزارهای توصیف مقایسه‌ای در این مطالعات، جدول (۱) در مقالات حاصل از این مطالعات است.

جدول (۱) در مقالات مورد شاهی :

- هدف آن مقایسه توزیع متغیرهای دموگرافیک و سایر متغیرهای مهم (بغیر از مواجهه) به تفکیک دو گروه مورد و شاهد، میباشد.
- در گزارش توصیف متغیرهای کمی از میانگین و انحراف معیار (یا میانه و صدکهای ۲۵ و ۷۵) و برای توصیف متغیرهای دسته‌بندی از تعداد و درصد، استفاده مینماییم.
- معنی‌داری آماری مقایسه هر یک از متغیرهای مورد نظر نیز در جدول (۱)، ارائه میگردد.
- دقت نمایید که بهتر است تحلیلگران مقالات، فقط نگاه انحصاری «معنی داری آماری» به تفاوتها نداشته باشند !!!! چرا؟؟؟

جدول (۱) در مقالات مورد شاهدهی : مثال یک مطالعه غیرهمسان

Factors associated with neonatal jaundice among neonates admitted to three hospitals in Burao, Somaliland: a facility-based unmatched case-control study

Yosef et al. *BMC Pediatrics* (2025) 25:710
<https://doi.org/10.1186/s12887-025-06036-2>

Table 1 Sociodemographic characteristics

Variables	Category	Case (64%)	Control (256%)	P-value
Mother age	18–25	15 (23.44%)	86 (33.59%)	0.294
	26–33	40 (62.50%)	138 (53.91%)	
	< 33	9 (14.06%)	32 (12.50%)	
Neonatal age	1–6	57 (89.06%)	225 (87.89%)	
	7–12	5 (7.81%)	14 (5.47%)	
	> 12	2 (3.13%)	17 (6.64%)	
Place of birth	Urban	48 (75.0%)	187 (73.05%)	0.752
	Rural	16 (25.0%)	69 (26.955)	
Occupation	Government	1 (1.565)	9 (3.52%)	0.576
	Merchant	6 (9.38%)	16 (6.25%)	
	Homeworker	52 (81.25%)	202 (78.91%)	
	Day labor	5 (7.81%)	29 (7.81%)	
Educational status	Illiterate	17 (26.56%)	75 (29.30%)	0.932
	Only read	13 (20.31%)	60 (23.44%)	
	Primary	15 (23.44%)	57 (22.27%)	
	Secondary	14 (21.88%)	46 (17.975)	
	College	5 (7.81%)	18 (7.03%)	
Sex of neonate	Male	41 (64.06%)	137 (53.52%)	0.129
	Female	23 (35.94%)	119 (46.48%)	



جدول (۱) در مقالات مورد شاهدهی : مثال یک مطالعه همسان

Smokeless Tobacco Usage and Oral Cancer Risk: A Hospital-Based Case-Control Study From Bangladesh

Tobacco Use Insights

Volume 18: 1-11

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1179173X251383752

journals.sagepub.com/home/tui



Abbreviation: ICD-10 International Classification of Diseases 10th Revision, BMI- Body Mass Index.

^aMean with Standard deviation.

^bp-value derived from chi-square test/Fisher-exact test.

Table 1. Descriptive Characteristics of the Study Population

Characteristics	Cases (%)	Controls (%)	P-value
Gender			—
Male	81 (47.9)	162 (47.9)	
Female	88 (52.1)	176 (52.1)	
Age			—
Mean age ^a	53.97 ± 12.47	53.96 ± 12.55	
Education			<0.001 ^b
No-formal education	98 (58.0)	141 (41.7)	
Primary school completed or less	47 (27.9)	100 (29.6)	
High school completed	18 (10.6)	62 (18.3)	
More than high school completed	6 (3.5)	35 (10.4)	
Work status			0.27 ^b
House worker	68 (40.2)	141 (41.7)	
Salaried worker	36 (21.3)	65 (19.2)	
Self-employed	28 (16.6)	50 (14.3)	
Farming	17 (10.1)	35 (10.4)	
Unemployed	20 (11.8)	47 (13.9)	
BMI (kg/m ²)			<0.001 ^b
Normal weight (18.50 -24.99)	68 (48.9)	91 (72.8)	
Under-weight (<18.50)	52 (37.4)	10 (8.0)	
Over-weight (≥25)	19 (13.7)	24 (19.2)	
Cancer site (ICD-10)			—
Tongue (C01-C02)	5 (3.0)	—	
Floor of the mouth (C04)	2 (1.2)	—	
Gum (C03)	20 (11.8)	—	
Soft palate(C05)	5 (3.0)	—	
Lower lip (C00.04)	2 (1.2)	—	
Cancer overlapping region (C06)	135 (79.8)	—	

آیا گزارش «معنی‌داری» در جدول (۱) به معنی تحلیل تک‌متغیره است؟

خیر! در واقع تحلیل تک‌متغیره، وقتی است که رابطه بین مواجهه /
مواجهات مورد نظر در هدف / اهداف مطالعه با پیامد مربوطه، مورد
ارزیابی قرار گیرد!

ملاحظات مهم و کاربردی جدول ۱ (۱)

• اول: بطور معمول سعی کنید **توزیع متغیر مواجهه / مواجهات مرتبط با اهداف مطالعه را**

در جدول (۱) نیاورید (منظور پرهیز از مخلوط شدن متغیرها).

• دوم: از نگاه تفسیری **مطلق** به شاخص معنی داری آماری یا P value، پرهیزید. به عبارت

دیگر و براساس منابع جدید، نگاه تفسیری که هر متغیر با تفاوت «معنی دار» بین دو گروه

مورد و شاهد، مخدوشگر / کوواریت مهم و تاثیرگذاری است و هر متغیر با تفاوت «غیر معنی

دار»، اینطور نیست، نگاه و رویکرد درستی نیست. رویکرد درست به شناسایی مخدوشگرها و

کوواریتها، نگاه مبتنی بر «دانش پیشین» و مرور متون است.

ملاحظات مهم و کاربردی جدول ۱ (۲)

- **سوم:** نتایج جدول (۱) مشابه جدول (۱) **کار آزماینها** تفسیر نمیشود (در کار آزماینها، جدول ۱ میتواند خود نشان از اجرای مطلوب راهکار «تصادفی سازی» باشد).
- **چهارم:** بهترین **رویکرد** در برخورد با **معنی داری آماری** یا **P value در جدول ۱**، این است که اگر متغیرهای دارای **«معنی داری آماری»**، در لیست مخدوشگرهای از پیش تعیین شده، نبود، میتوانید آنها را علاوه بر متغیرهای مخدوشگر / کوواریت پیشین، در یک رویکرد **تحلیل حساسیت**، کنترل یا تصحیح نمایید.
- **پنجم:** بهتر است در جدول (۱)، به پدیده **Missing data** در متغیرهای مورد نظر نیز توجه

چگونه جدول ۱ را برای مقالات مورد نظر، آماده کنیم؟

• راهکار اول: آماده سازی یک **تمپلیت در نرم افزار Word** و سپس مقایسه توزیع دو دسته

متغیرهای کمی و دسته بندی و اجرای آزمونهای آماری مربوطه و در انتها، وارد نمودن

یافته‌های فوق در تمپلیت فوق

• راهکار دوم: استفاده از **ماژول table1_mc در Stata version 16** یا بالا تر، بمنظور

آماده سازی اتوماتیک این جدول برای مقاله

کار عملی (۱): آماده سازی جدول ۱ در یک مطالعه مورد شاهدی

ابتدا در دیتاست NTD and Zinc.dta با هدف اصلی «تعیین رابطه بین سطح سرمی روی در مادران باردار و وقوع ناهنجاری نقص لوله عصبی در نوزادان آنها» که حاصل یک مطالعه مورد شاهدی غیرهمسان است، متغیرهای مناسب برای جدول ۱ را انتخاب نموده و با استفاده از مازول table1_mc، جدول مربوطه را آماده نموده و آنرا تفسیر نمایید.

انواع تحلیلهای تک متغیره در مطالعات مورد شاهدهی :

- تحلیل تک متغیره با جدول 2×2 (مواجهات دو سطحی یا باینری)
- تحلیل تک متغیره با جدول $r \times 2$ (مواجهات چند سطحی یا دسته بندی)
- تحلیل تک متغیره برای مواجهات کمی

تحلیل تک متغیره با جدول 2×2 : Unmatched CCS

	Case	Control	Total
Exposure +	a	b	a+b
Exposure -	c	d	c+d
Total	a+c	b+d	n

✓ ساده ترین رویکرد تحلیل تک متغیره در مطالعات مورد شاهدهی است.

✓ به ازای هر جدول، دو شاخص فراوانی (فراوانی مواجهه در دو گروه مورد و شاهد)، یک آزمون کای اسکویر و یک شاخص OR خواهیم داشت.

✓ در استیتا با استفاده از دستور “cc” میتوانیم این تحلیل را در مطالعات مورد شاهدهی غیرهمسان، برآورد نماییم.

✓ فرمت دیتا در مطالعات مورد شاهدهی غیرهمسان، باید Long format باشد.

✓ هر دو متغیر پیامد و مواجهه باید دارای کدهای «صفر» و «یک» باشد.

تحلیل تک متغیره با جدول 2×2 : Matched CCS

	Exp+ (cont)	Exp- (cont)	Total
Exp + (case)	a	b	a+b
Exp - (case)	c	d	c+d
Total	a+c	b+d	n

✓ فرمت جدول در این مطالعات، تغییر مینماید.

✓ به ازای هر جدول، دو شاخص فراوانی (فراوانی مواجهه در دو گروه مورد و شاهد)، یک آزمون مک نمار و یک

شاخص OR، تحت عنوان matched OR خواهیم داشت!

✓ در استیتا با استفاده از دستور "mcc" میتوانیم این تحلیل را در مطالعات مورد شاهدهی همسان، برآورد نماییم.

✓ فرمت دیتا در مطالعات مورد شاهدهی غیرهمسان، باید Wide format باشد.

✓ هر دو متغیر پیامد و مواجهه باید دارای کدهای «صفر» و «یک» باشد.

تحلیل تک متغیره با جدول $2 \times r$: Unmatched CCS

	Case	Control	Total
Exp (L2)	a	b	a+b
Exp (L1)	c	d	c+d
Exp (ref)	e	f	e+f
Total	a+c+...	b+d+...	n

✓ برای مواجهات ۳ سطح و بیشتر استفاده میشود

✓ به ازای هر جدول، معمولاً یک آزمون کای اسکویر (کلی) و بیش از یک

شاخص OR خواهیم داشت (به تعداد لایه های منهای یک شاخص OR). در صورتیکه متغیر مواجهه، رتبه ای باشد، میتوان از یک آزمون «کای اسکویر برای روند» نیز استفاده نمود.

✓ در استیتا با استفاده از دستور “**tabodds**” میتوانیم این تحلیل را در مطالعات مورد شاهدهی غیرهمسان، برآورد نماییم.

✓ فرمت دیتا در مطالعات مورد شاهدهی غیرهمسان، باید Long format باشد.

✓ متغیر مواجهه باید دارای کدهای «صفر» (لایه مرجع)، «یک» (لایه بعدی)، «دو» (لایه بعدی) و ... باشد.

تحلیل تک متغیره با جدول $2 \times r$: Matched CCS

✓ برای مواجهات ۳ سطح و بیشتر در طراحیهای Matched، دیگر جدول مشابه موارد قبلی، مورد استفاده قرار نمیگیرد.

✓ در این موارد، تنها روش مورد استفاده، روش «رگرسیون لجستیک شرطی یا Conditional Logistic **Reg**» میتوانیم استفاده نماییم.

✓ در استیتا با استفاده از دستور “**clogit**” میتوانیم این تحلیل را در مطالعات مورد شاهدهی همسان، برآورد نماییم.

✓ فرمت دیتا در این مطالعات مورد شاهدهی همسان، باید Long format باشد !!!

✓ متغیر مواجهه باید دارای کدهای «صفر» (لایه مرجع)، «یک» (لایه بعدی)، «دو» (لایه بعدی) و ... باشد.

✓ متغیر همسان یا matching variable نیز ضروری است.

تحلیل تک متغیره با مواجهه کمی : Unmatched/ Matched CCS

✓ در این موارد نیز دیگر از جداول استفاده نمیشود.

✓ در این موارد، برای آزمون آماری از «آزمون تی مستقل (برای غیرهمسان شده) یا تی زوج (برای همسان شده)» باید استفاده نماییم.

✓ در استیتا با استفاده از دستور “ttest” میتوانیم این تحلیل را در مطالعات مورد شاهدهی غیرهمسان یا همسان، برآورد نماییم.

✓ فرمت دیتا در این مطالعات مورد شاهدهی غیرهمسان، Long format و در مورد همسان، باید Wide format باشد !!!

✓ بعنوان شاخصهای اندازه اثر نیز از شاخصهای MD / SMD استفاده میگردد.

کار عملی (۲): تحلیل تک متغیره در یک مطالعه مورد شاهدی غیر همسان

✓ ابتدا در دیتاست LBW_unmatched CCS.dta که یک مطالعه

غیر همسان با نسبت تقریبی شاهد به مورد ۲ به ۱ است، برای هدف اصلی

«رابطه بین وزن مادران در ابتدای بارداری با تولد نوزاد کم وزن» تحلیل مناسب را

انجام داده و نتیجه را تفسیر نمایید.

✓ چنانچه در همین دیتا، بخواهیم هدف دوم «رابطه بین مصرف سیگار در دوران

بارداری و حتی قبل از بارداری با وقوع تولد نوزاد کم وزن» را ارزیابی نماییم،

چگونه باید تحلیل انجام شود؟ تفسیر چیست؟

کار عملی (۳): تحلیل تک متغیره در یک مطالعه مورد شاهدی همسان 1:1

✓ ابتدا در دیتاست LBW2_matched CCS.dta که یک مطالعه همسان با نسبت شاهد به مورد ۱ به ۱ است، برای هدف اصلی «رابطه بین وزن مادران در ابتدای بارداری با تولد نوزاد کم وزن» تحلیل مناسب را انجام داده و نتیجه را تفسیر نمایید.

✓ چنانچه در همین دیتا، بخواهیم هدف دوم «رابطه بین مصرف سیگار در دوران بارداری و حتی قبل از بارداری با وقوع تولد نوزاد کم وزن» را ارزیابی نماییم، چگونه باید تحلیل انجام شود؟ تفسیر چیست؟

مشاهده لیست دوره‌های آموزشی www.researchware.org/shop

پایان درس سوم
این دوره ادامه دارد ... !

برای کسب اطلاع از آخرین اخبار آموزشی کلیه دوره های آموزشی آکادمی پژوهش‌افزار (دکتر کشتکار)،
در گروه تلگرامی زیر عضو شوید: @amarafzar2011