

# روش تحقیق و آمار زیستی

## درسنامه‌ی تشریحی

گردآورندگان 

یوسف مرادی

|                     |   |
|---------------------|---|
| شابک                | 978-600-7827-91-8   |
| شماره کتابشناسی ملی | ۴۱۸۱۸۰۱   |
| عنوان و نام پدیدآور | روش تحقیق و آمار زیستی / گردآورنده یوسف مرادی.                |
| مشخصات نشر          | تهران: انتشارات علمی سنا                                      |
| مشخصات ظاهری        | ۲۹۴ ص.: مصور، جدول، نمودار (بخشی رنگی).                       |
| یادداشت             | کتابنامه: ص. ۲۹۴.   |
| موضوع               | زیست‌سنجی -- تحقیق -- روش‌شناسی -- راهنمای آموزشی (عالی)      |
| موضوع               | زیست‌سنجی -- تحقیق -- روش‌شناسی -- آزمون‌ها و تمرین‌ها (عالی) |
| موضوع               | تحقیق -- روش‌شناسی -- راهنمای آموزشی (عالی)                   |
| موضوع               | تحقیق -- روش‌شناسی -- آزمون‌ها و تمرین‌ها (عالی)              |
| رده بندی دیویی      | ۱۵۱۹۵۰۷۲/۵۷۰  |
| رده بندی کنگره      | ۴۹۳۲۴QH / م   |
| سرشناسه             | مرادی، یوسف، ۱۳۶۸ - گردآورنده                                 |
| وضعیت فهرست نویسی   | فیبا  |



**موسسه علمی انتشاراتی سنا**

روش تحقیق و آمار زیستی

یوسف مرادی

۹۷۸-۶۰۰-۷۸۲۷-۹۱-۸

دوم - ۱۴۰۳

انتشارات علمی سنا

لیلا انصاری

elmisana@gmail.com

sanabook.com

۱۰۰ نسخه

برای مشاهده قیمت اسکن کنید. ➔

نام کتاب

گردآورندگان

شابک

سال چاپ

صفحه‌آرایی

طراح جلد

پست الکترونیک

فروش اینترنتی

تیراژ

قیمت

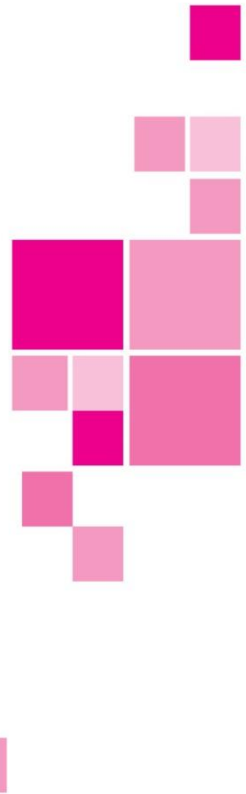


## فهرست :

- بخش اول : آمار و احتمالات ..... ۱
- فصل اول: توصیف عددی ..... ۱
- فصل دوم: توزیع های فراوانی و نمودارها ..... ۹
- فصل سوم: شاخص های مرکزی ..... ۲۹
- فصل چهارم: شاخص های پراکندگی ..... ۴۵
- فصل پنجم: نمرات استاندارد ..... ۶۲
- فصل ششم: احتمال و توزیع های آن ..... ۷۲
- فصل هفتم : همبستگی ..... ۱۰۲
- فصل هشتم: رگرسیون و پیش بینی ..... ۱۱۶
- فصل نهم: نمونه گیری و آزمون فرضیه ..... ۱۳۷
- فصل دهم: آزمون های پارامتریک و ناپارامتریک ..... ۱۶۱
- فصل یازدهم: مقایسه های چندگانه میانگین های مورد آزمایش آزمونهای تعقیبی ..... ۲۰۱
- بخش دوم : روش تحقیق ..... ۲۰۷
- فصل اول : تحقیق و مفاهیم اساسی آن ..... ۲۰۷
- فصل دوم: خطا در پژوهش ..... ۲۱۶
- فصل سوم : انواع مطالعات ..... ۲۲۲
- فصل چهارم: جامعه و نمونه ..... ۲۲۶
- فصل پنجم: تحقیق زمینه یابی ..... ۲۳۲

|     |  |
|-----|--|
| ۲۴۳ | فصل ششم: تحقیق آزمایشی .....   |
| ۲۵۱ | فصل هفتم: آزمایش میدانی و طرح های درون آزمودنی ها .....              |
| ۲۵۸ | فصل هشتم: پژوهش تاریخی .....   |
| ۲۶۴ | فصل نهم: پژوهش در عمل .....  |
| ۲۶۹ | فصل دهم: تحقیق پس رویدادی .....                                      |
| ۲۷۷ | فصل یازدهم : تحقیق کیفی .....  |
| ۲۸۹ | سوالات تستی سال های گذشته کنکور کارشناسی ارشد روانشناسی بالینی ..... |
| ۲۹۷ | پاسخ های تشریحی .....  |

بخش اول : آمار و احتمالات  
فصل اول:  
توصیف عددی





## بخش اول : آمار و احتمالات

### فصل اول: توصیف عددی

#### هدف آمار:

روشهای آماری دارای دو وظیفه اند:

۱- آمار توصیفی (گردآوری و سازماندهی داده ها)

۲- آمار استنباطی (استخراج و استنباط اطلاعات)

جامعه: گروهی از افراد ، اشیا یا حوادث که حداقل دارای یک صفت یا ویژگی مشترک هستند. ماهیت پژوهش تعیین کننده دامنه جامعه است.

پارامتر: ویژگی عددی یک جامعه است.

آماره : ویژگی یا ویژگی های کمی که یک نمونه را توصیف می کنند. آماره پارامتر را برآورد می کند.

معرفی بودن نمونه غالباً از طریق انتخاب نمونه بصورت تصادفی امکان پذیر می باشد.

متغیر: ویژگیهایی که بیش از یک ارزش به آنها اختصاص داده می شود و تغییرات را از فردی به فرد دیگر یا از شی به شی دیگر نشان میدهد.

#### انواع متغیر:

۱. متغیرهای کمی و کیفی ← متغیرهای کمی ← قابل اندازه گیری می باشند (قد، وزن، فشارخون و...)  
متغیرهای کیفی ← غیر قابل اندازه گیری (رنگ چشم، جنس، ...)

کمی گسسته:  
کمی پیوسته:  
کمی

۲. متغیرهای مستقل وابسته



۷ - گزینه (د) سطح سوال متوسط

درصد اشغال تخت یک متغیر نسبتی است چون دارای صفر ذاتی می باشد. یعنی میتوان گفت به فرض مثال درصد اشغال در منطقه ای صفر است.

۸ - گزینه (الف) سطح سوال آسان

چون کیفی است بنابراین یا اسمی است یا رتبه ای از طرف دیگر چون قابل رتبه بندی نیستند مثل طبقه اجتماعی یا بیماری بنابراین اسمی است.

۹ - گزینه (الف) سطح سوال متوسط

ترتیب مقیاس ها به صورت زیر است: اسمی رتبه ای فاصله ای نسبتی مقیاس های پایین تر دارای ویژگی های مقیاس بالاتر به اضافه یک ویژگی جدید است و بنابراین مقیاس نسبتی علاوه بر داشتن یک ویژگی جدید تمام ویژگی های مقیاس های بالاتر را دارد.

۱۰ - گزینه (ب) سطح سوال متوسط

متغیرها دو دسته هستند: ۱- کمی (فاصله ای و نسبتی) ۲- کیفی (اسمی و ترتیبی (رتبه ای))

۱۱ - گزینه (ب) سطح سوال آسان

۱۲ - گزینه (ج) سطح سوال آسان

۱۳ - گزینه (ب) سطح سوال آسان

اولا چون عدد است پس کمی است و از طرف دیگر گفته به صورت سال تحصیلی (نیمسال) بنابراین گسسته است.

۱۴ - گزینه (ج) سطح سوال متوسط

چون متغیرهای فاصله ای دارای صفر ذاتی نمی باشند بنابراین نسبت اندازه ی آن صفت غیرقابل تشخیص می باشد.

۱۵ - گزینه (ج) سطح سوال آسان

۱۶ - گزینه (ج) سطح سوال آسان

۱۷ - گزینه (ب) سطح سوال متوسط



مقیاس نسبتی بالاترین مقیاس است و دارای صفر ذاتی می باشد.

۲۷- گزینه (ج) سطح سوال آسان

چون دما یا درجه ی حرارت فاقد صفر ذاتی می باشد.

۲۸- گزینه (ج) سطح سوال آسان

۲۹- گزینه (الف) سطح سوال متوسط

هر دو متغیر دارای صفر ذاتی هستند بنابراین یک متغیر نسبتی می باشند.

۳۰- گزینه (ب) سطح سوال متوسط

چون زایمان تحت تاثیر تمرینات ورزشی قرار می گیرد بنابراین طول زایمان یک متغیر مستقل است.

۳۱- گزینه (الف) سطح سوال آسان

۳۲- گزینه (ج) سطح سوال آسان

۳۳- گزینه (د) سطح سوال متوسط

چون فرهنگ جامعه را نامگذاری کرده است و هر کدام در یک دسته قرار دارند که هیچ کدام نسبت به دیگری برتری ندارند.

۳۴- گزینه (ب) سطح سوال متوسط

چون صدا باعث حوادث ناشی از کار می شود.

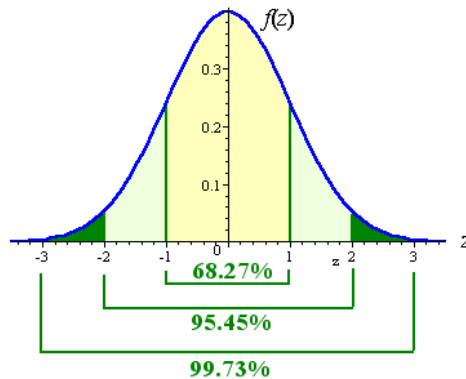
۳۵- گزینه (د) سطح سوال متوسط

چون این متغیر کمی بوده و در ضمن دارای صفر ذاتی است یعنی صفر دندان پوسیده به طور مطلق صفر بوده و با معنی است این متغیر نسبتی می باشد.

۳۶- گزینه (د) سطح سوال دشوار

با توجه به اینکه تعداد دخترها و پسرها مشخص می باشد لذا با یک جدول دوطرفه که در یک طرف آن جنس و در طرف دیگر وجود یا عدم وجود مشکل ژنتیکی است؛ می توان داده ها را خلاصه کرد. ولی اگر در نمونه ۱۷ نفری تعداد دخترها و پسرها نامشخص باشد تنها می توان از جدول یک طرفه یا نمودار ستونی استفاده کرد.





محدودیت  $z$ : ۱. منفی می شود ۲. اعشاری می شود.

نمره استاندارد  $t$  اگر نمره  $z$  را ضرب در ۱۰ به اضافه ۵۰ کنیم  $t$  به دست می آید.

$$t = z \times 10 + 50$$

در توزیع میانگین ۵۰ و انحراف معیار ۱۰ است. بین ۲۰ تا ۸۰ -

$$IO = Z \times 15 + 100$$

$IQ =$  یکی دیگر از نمره های استاندارد است

$CEEB =$  شاخص استاندارد پیشرفت تحصیلی

$$CEEB = z \times 100 + 500$$

$$a_z = \frac{a - \mu}{\delta}$$

فرمول استاندارد توزیع نرمال:

مثال: قد ساکنین منطقه ای از توزیع نرمال با میانگین ۱۶۵ و واریانس ۸۱ پیروی می کند. اگر فردی را به

طور تصادفی از میان ساکنین این منطقه انتخاب کنیم احتمال اینکه: قد وی کمتر از ۱۶۰ باشد چقدر است؟

$$x = \text{قد}$$

$$a_z = \frac{a - \mu}{\delta} = \frac{160 - 165}{9} = 0/56$$

نکته) در کلیه توزیع های نرمال سطح زیر منحنی در فاصله  $\mu \pm 1\delta$  تقریباً برابر ۰/۶۸۲۶ و در فاصله

$\mu \pm 2\delta$  تقریباً برابر ۰/۹۵۴۴ و در فاصله  $\mu \pm 3\delta$  برابر ۰/۹۹۷۴ است.

نکته) در علوم پزشکی و بهداشت  $\mu \pm 2\delta$  را به عنوان حد نرمال یک اندازه می شناسند.



## فصل هشتم: رگرسیون و پیش بینی

برای بررسی ارتباط بین صفات کمی از دو روش استفاده می شود ← (۱) همبستگی Correlation

(۲) رگرسیون (Regression)

رگرسیون (Regression) ⇐ برای بیان رابطه ای کاربردی میان دو متغیر به کار برده می شود به طوری که مقدار یک متغیر را بتوان با استفاده از متغیر دیگر پیش بینی کرد.

برآورد ضریب رگرسیون: در مدل خطی  $y = a + bx$  که در آن  $a$  و  $b$  ضرایب رگرسیون نامیده می شوند.  $b$  ضریب زاویه و  $a$  ضریب پایه نام دارند، در واقع این ضریب پارامترهایی از جامعه می باشند. برای یافتن برآورد از این خط باید  $a$  و  $b$  را برآورد نمود. برای برآورد خطی این پارامترها بایستی چنان محاسبه شوند که مجموع مجذورات فاصله نقاط تا خط رگرسیون یا به عبارت دیگر مجموع مجذورات فاصله  $y$  های مشاهده شده از  $y$  های بدست آمده توسط خط رگرسیون حداقل شود. بر این اساس مقدار  $b$  عبارتست از:

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{SP_{xy}}{SS_x}$$

$$a = \bar{y} - b\bar{x}$$

ضریب همبستگی نشان دهنده ارتباط دو متغیر است اما زمانی که بدنبال اثر یک متغیر بر متغیر دیگر هستیم نمی توان از آن استفاده کرد. از طرفی گاهی تخمین تغییر در یک متغیر با تغییر متغیر دیگر برای ما مهم است که با هم ضریب همبستگی نمی تواند تخمینی از این تغییرات را ارائه دهد. برای پاسخ به سوالات بالا باید سراغ رگرسیون برویم.

در رگرسیون یک متغیر که تاثیرپذیر ( $y$ ) از سایر متغیرهاست، متغیر پاسخ (وابسته) نامیده می شود و متغیرهایی که بر متغیر پاسخ اثر می گذارند ( $x$ ) متغیر مستقل نامیده می شوند.

**معادله خط رگرسیون:** بین دو متغیر  $x$  و  $y$  همیشه باید رابطه خطی باشد اگر در یک متغیر فاصله ای حداقل ۳۰ نفر بگیریم و نمودارشان را رسم کنیم. رابطه خطی می بینیم.



الف) ۰/۱۶ (ب) ۰/۰۸ (ج) ۰/۰۴ (د) ۰/۰۲

۴۱- از یک توزیع نرمال با میانگین و انحراف معیار ۱۰ نمونه بی تصادفی به حجم ۲۵ انتخاب می کنیم. احتمال اینکه میانگین این نمونه کمتر از ۱۴ باشد تقریباً چقدر است؟

الف) ۰/۹۷۵ (ب) ۰/۹۹ (ج) ۰/۹۰ (د) ۰/۹۵

۴۲- از یک جامعه که فقط شامل ۹ فرد است و انحراف معیار قد آن ها ۶ سانتیمتر است تمام نمونه های ممکن ۴ تایی (با جایگذاری) انتخاب و میانگین هر نمونه به دست آمده است انحراف معیار میانگین های حاصل عبارت است از:

الف) ۳ (ب) ۶ (ج) ۲ (د) ۱/۵

۴۳- اگر متغیر تصادفی  $X$  دارای توزیع نرمال با میانگین ۱۲ و واریانس ۲۷ باشد. میانگین نمونه بی  $\bar{X}$  که براساس نمونه گیری تصادفی به حجم ۹ به دست می آید دارای کدام توزیع و به ترتیب با چه مقدار میانگین و انحراف معیار می باشد؟

الف)  $t$  با میانگین ۱۲ و انحراف معیار  $\sqrt{3}$  (ب)  $t$  با میانگین ۱۲ و انحراف معیار ۳  
ج) نرمال با میانگین ۴ و انحراف معیار ۳ (د) نرمال با میانگین ۱۲ و انحراف معیار  $\sqrt{3}$

۴۴- اگر انحراف معیار صفت  $X$  در یک جامعه برابر ۲۰ باشد. حجم نمونه لازم برای آنکه با ۹۵ درصد اطمینان بتوان گفت خطای برآورد میانگین کمتر از ۵ واحد است تقریباً برابر است با:

الف) ۶۲ (ب) ۸ (ج) ۱۵۷ (د) ۲۴۷

پاسخ های تشریحی :

سطح سوال دشوار

۱-گزینه (ب)



نکته: در صورت مقایسه متغیر کیفی چند حالت در صورت نرمال بودن جامعه از آنالیز واریانس استفاده میشود. در صورت نرمال نبودن جامعه یا شرایط غیر پارامتریک از کروسکال والیس استفاده میکنیم.

### ۱۹. گزینه (الف) سطح سوال متوسط

درجه ی آزادی خطا که همان درجه ی آزادی درون نمونه ها است عبارت است از:  $(n - k)$  (که  $n$  تعداد

کل افراد در همه ی نمونه ها است و  $n = \sum_{i=1}^k n_i$  و  $k$  تعداد نمونه ها است). بنابراین

$$d.f = \sum n_i - k$$

برای مسأله مذکور می توان نوشت

### ۲۰. گزینه (ب) سطح سوال دشوار

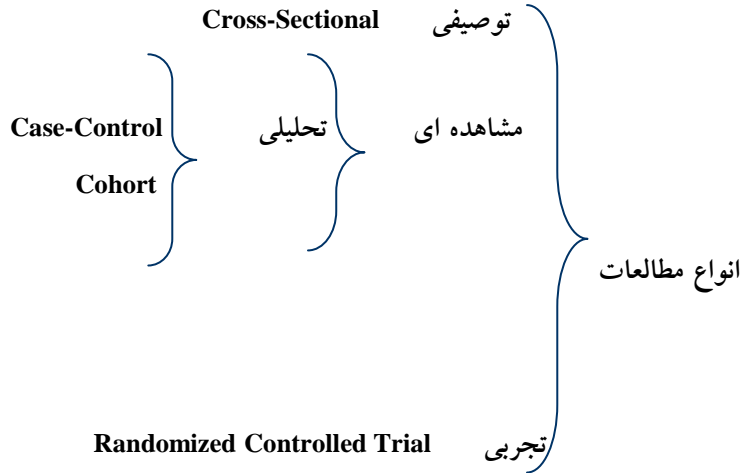
آنالیز واریانس با استفاده از واریانس (یک اندازه پراکندگی) تفاوت بین میانگین ها (که یک اندازه گرایش مرکزی است) را بررسی می کند. در آنالیز واریانس دو نوع واریانس مورد تحلیل قرار می گیرد. واریانس بین گروهی که مقدار تفاوت میانگین های نمونه را از یکدیگر اندازه گیری می کند. هر قدر تفاوت میانگین ها بیشتر باشد واریانس بین گروه ها بیشتر است. واریانس نوع دوم واریانس درون گروهی نام دارد و پراکندگی داده های درون هر گروه را اندازه گیری می کند. دو اصل برای تعیین معنی دار بودن تفاوت میانگین با استفاده از واریانس وجود دارد. ابتدا هر قدر واریانس بین گروهی بزرگتر باشد احتمال وجود تفاوت معنی دار بیشتر است. دوم هر قدر واریانس درون گروهی کوچکتر باشد احتمال وجود تفاوت معنی دار بیشتر است. آماره تحلیل واریانس و نمره  $F$  این دو اصل را با محاسبه نسبت این دو نوع واریانس ترکیب می کند.

$$f = \frac{\text{انس بین گروهی}}{\text{انس درون گروهی}}$$

اگر دو اصل مورد بحث را در نظر بگیریم بدین مفهوم است که یک نمره  $F$  بالاتر احتمال وجود تفاوت معنی دار بین میانگین ها را افزایش می دهد. بنابراین احتمال درست بودن فرض صفر در این حالت کمتر خواهد بود. زمانی که نمره  $F$  محاسبه شده مساوی یا بیشتر از مقدار جدول باشد (مقدار  $P$  مساوی یا کمتر از مقدار از پیش تعیین شده خواهد بود). در این شرایط فرض صفر رد می شود. در مسأله ی مذکور میانگین مربعات بین گروه ها و داخل گروه ها همان واریانس بین گروه ها و داخل گروه ها است. وقتی



## فصل سوم : انواع مطالعات



## مطالعه مقطعی (cross sectional)

در یک مطالعه مقطعی، پژوهشگر تمام سنجش هایش را در یک زمان خاص انجام می‌دهد. وی نمونه‌ای از جمعیت را می‌گیرد و به توزیع متغیرها در داخل آن نمونه نگاه می‌کند، سپس ممکن است از روابط بین متغیرهایی که تصمیم می‌گیرد (با استفاده از اطلاعات منابع گوناگون) به عنوان مستقل و وابسته طراحی نماید استنتاج علت و معلولی کند. در یک مطالعه مورد - شاهدی، وارونه عمل می‌کند. وی با پیامد، شروع می‌کند، نمونه‌ای از جمعیت بیماران مبتلا (موارد) و نمونه دیگری از جمعیت بدون آن بیماری (شاهد‌ها) را انتخاب می‌کند، سپس سطوح متغیرهای مستقل را در دو نمونه با هم مقایسه می‌نماید تا ببیند کدامیک با پیامد بیماری رابطه دارند.

یک نقطه قوت عمده مطالعه‌های مقطعی بر مطالعه‌های همگروهی (و مطالعه‌های تجربی) این است که برای رویداد پیامد نباید انتظار کشید. این امر آن‌ها را سریع و ارزان نموده و بدین معنی است که مسئله گم شدن افراد در حین پیگیری وجود ندارد. مطالعه مقطعی را می‌توان با هزینه اندک یا بدون هزینه اضافی به عنوان نخستین مرحله یک مطالعه همگروهی یا تجربی در نظر گرفت. نتایج آن ویژگی‌های جمعیتی و