Date: Subject:

تعریف آمار: با استفاده از علم آمار می‌توان به جمع آوری داده های مورد نیاز به منطق رسیدن به اطلاعات لازم با استفاده از درصد، فراوانی، جدول، نمودار دست یافت

آمار ← توصیفی
آمار ← استنباطی

انواع آمار: جمع آوری داده و تحقیق آن را در علم آمار می‌توان در قالب آمار توصیفی و آمار استنباطی ارائه نمود.

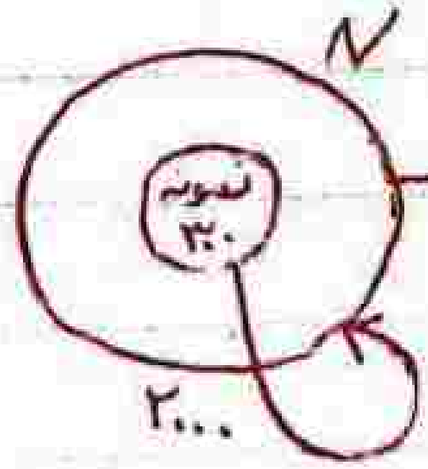
آمار توصیفی: در آمار توصیفی هرآنچه که هست بدون هیچ گونه کم و کاست در قالب فراوانی، درصد و جدول و نمودار ارائه می‌گردد. مانند توصیف شرایط یک کلاس درس یا همانند سرشماری نفوس و مسکن

آمار استنباطی: نتایج حاصل از یک نمونه کوچک به جامعه بزرگتر تعمیم داده می‌شود در آمار استنباطی نمونه گیری صورت می پذیرد و نتایج نمونه ها مورد بررسی قرار می‌گیرد

جامعه (تعداد) N (تعداد) نمونه n

جامعه و نمونه: جامعه به مجموعه ای از افراد و اشیا اطلاق می‌شود که حداقل دارای یک صفت مشترک می باشد و با N نشان داده می‌شود مانند جامعه حسابرسان ایران جامعه فرشتگان

اما نمونه زیر جامعه‌ای است که از داخل جامعه انتخاب شده و معرف ویژگی های جامعه می باشد



Scanned with CamScanner

Date Subject

متغیر و انواع آن:

متغیر در معنای لغوی یعنی هرچیزی که تغییر نماید اما در معنای عملیاتی به پدیده‌ای که مورد بررسی قرار می‌گیرد و متغیر اطلاق می شود

الف) متغیر برمبنای نحوه جمع آوری داده

هرگاه متغیری برمبنای داده های عددی جمع آوری شود به آن متغیر کمی می گویند مانند سن، وزن، حجم

و متغیری که با داده های عددی جمع آوری نشود متغیر کیفی نامیده می‌شود. مانند جنسیت نوع حوزه، پست سازمانی، نوع شرکت

متغیر کمی که قابلیت پذیرفته شدن اعشار را داشته باشد کمی پیوسته نامیده می‌شود مانند وزن، حجم، مسافت و متغیر کمی که قابلیت پذیرفته شدن اعشار نداشته باشد متغیر کمی گسسته نامیده می شود مانند تعداد دانشجویان، تعداد شرکت، تعداد صندلی

ب) متغیرها برمبنای نقش و صرف

متغیرها برمبنای نقشی که در ترکیب فرآیند تحلیلی دارند به ۲ متغیر اصلی، مستقل و وابسته تقسیم می‌شود

متغیر مستقل: متغیری است که در اختیار پژوهشگر بوده قابل دستکاری بوده و تاثیر این دستکاری بر متغیر وابسته ظاهر می‌شود

اما متغیر وابسته تاثیر پذیر بوده و در اختیار پژوهشگر نمی باشد

نکته: هر متغیری که قبل از حروف ربط (با - بر - و - در) بیاید متغیر مستقل بوده و متغیرهای بعد از این حروف نیز وابسته محسوب می شود

تاثیر جنسیت بر عملکرد کاری دبیران

مستقل وابسته

تاثیر هوش بر معدل

مستقل وابسته

Scanned with CamScanner

Date: Subject:

مقیاس اندازه	نوع متغیر	نوع تحلیل	ویژگی
۱- اسمی	کیفی	ناپارامتریک	طبقه‌بندی، نام‌گذاری - دسته‌بندی
۲- رتبه‌ای			طبقه‌بندی - نام‌گذاری - دسته‌بندی - نظم در طبقات
۳- فاصله‌ای	کمی	پارامتریک	صفر قراردادی یعنی عدد ساختگی
۴- نسبی			صفر مطلق یعنی عدد واقعی

اگر در بین گروه‌ها نظم وجود داشت (یعنی تر و ترین داریم بین بد و بدتر و بهتر و خوب‌تر) رتبه‌ای است اگر نظم وجود نداشت اسمی است

قوی‌ترین ضعیف‌ترین

هرچه مقیاس ما قوی‌تر باشد عمومیت پیدا می‌کند و تعمیم‌پذیری پیدا می‌کند

آمار ناپارامتریک مثل شربت ایرانی همان ضعیفه است و مقیاس و داده آن ضعیف و کیفی است

آمار پارامتریک مثل شربت خارجی همان قویه است

هوش مردان	هوش زنان
۱۰۵	۱۴۰
۱۲۰	۱۱۵
۱۲۴	۱۴۵
۱۳۵	۱۰۰

پارامتریک ← مستقل ← t

ناپارامتریک ← U

مقیاس نسبی داده‌ها کمی و پارامتریک است و از صفر مطلق پیروی می‌کند چون عدد آن واقعی است. حساب ما صفر است یعنی دلیل بر تهی بودن دارد. اما اگر نمره من صفر است دلیل بر تهی بودن ندارد شرایط ایجاب کرده که صفر بگیرد.

تصادف با سرعت ۱۲۰ $\frac{km}{h}$ است از کدام مقیاس پیروی می‌کند؟ کمی است و چون عدد واقعی است پس نسبی است

Scanned with CamScanner

Date: Subject:

برای نمره من ۱۵ است و کدی است علت چیست؟ آیا در ایران بین ۵ تا ۱۵ فقط نمره است در کشورهای دیگر چگونه است. چون نمره برحسب قرارداد است بین فاصله‌ای است.

و آزمون: درس، میانگین ← فاصله‌ای
لیتر، حجم، سرعت ← نسبی

مقیاس یا ماسیک که متغیر از ۴ حالت خارج نمی‌باشد

الف) مقیاس اسمی: این مقیاس ضعیف‌ترین مقیاس اندازه‌گیری محسوب می‌شود و در این مقیاس می‌توان افراد و اشیا را دسته‌بندی و طبقه‌بندی و نام‌گذاری و کد‌گذاری نمود. اعداد به‌کار رفته در این مقیاس معنا و مفهوم کمی ندارند و صرفا جهت تفکیک طبقات استفاده می‌شود. مناسب‌ترین شاخص آماری برای این مقیاس مد یا نما می‌باشد بیشترین تکرار مد است.

ب) مقیاس رتبه‌ای: این مقیاس تمام ویژگی‌های مقیاس اسمی را دارا است با این تفاوت که می‌توان در این مقیاس ترتیب قرارگرفتن افراد و اشیا را برحسب یک ملاک از پیش تعیین شده مشخص نمود. نکته کلیدی این مقیاس تر می‌باشد. مناسب‌ترین شاخص آماری در این مقیاس میانه می‌باشد (نقطه ۵۰ درصد)

ج) مقیاس فاصله‌ای: این مقیاس علارغم ۲ مقیاس قبلی با اعداد سروکار داشته و اعداد به کار رفته در این مقیاس به دلیل وجود صفر قراردادی اعداد ساختگی می‌باشد. مناسب‌ترین شاخص آماری در این مقیاس میانگین است.

نکته: هرعددی که با استفاده از امتحان، آزمون، پرسشنامه، بدست آید از مقیاس فاصله‌ای پیروی می‌کند.

Bian SHAD

Scanned with CamScanner

Date Subject

د) مقیاس نسبی: این مقیاس کامل ترین مقیاس اندازه گیری بوده و به دلیل برخورداری از صفر مطلق اعداد آن واقعی می باشد مانند حجم، مسافت، وزن، و مناسب ترین شاخص آماری آن انحراف معیار است می باشد
قوی ترین شاخص اندازه گیری

X = داده خام

$\bar{X}$ = میانگین $\Longrightarrow x = X - \bar{X}$ تفاضل عدد خام و میانگین

C یا i = فاصله طبقات

cl = طبقات

D یا R = دامنه تغییرات

S^2 یا G^2 = واریانس

واریانس: پراکندگی هر عدد را از میانگین محاسبه می کند مثلا نمرات ۲ گروه A و B را داریم که میزان تفاوت و پراکندگی گروه B کمتر و میزان تفاوت و پراکندگی در گروه A بیشتر است که گروه A واریانس است

چون همه اعداد در آن وجود دارد پس این واریانس است

A: ۱۶_۱۵_۲۰_۱۸_۱۹_۱۴

B: ۱۷_۲۰_۲۰_۱۹_۱۸_۱۹

N یا n = تعداد (تعداد نمونه، تعداد جامعه)

F یا f = فراوانی

Q = چارک (نقطه ۲۵ درصد)

Q_1 = چارک اول Q_2 = میانه یا Md (۵۰ درصد)

Q_3 = چارک سوم (نقطه ۷۵ درصد)

S یا G = انحراف استاندارد p = درصد $\sum$ = مجموع

r = همبستگی
correlation

cF یا Fc = فراوانی تجمعی - فراوانی تراکمی

$$\sum XY$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

Scanned with CamScanner

هرگونه محاسبه در آمار با یک شاخص معرفی می شود که به شاخص های آماری معروف می باشند کلیه شاخص های آماری را می توان در ۲ دسته شاخص های مرکزی و شاخص های پراکندگی قرار داد. منظور از شاخص مرکزی، شاخصی است که احتمالا در مرکز اعداد یافت می شود اصطلاحا گرایش به مرکز توزیع فراوانی دارد که معروف ترین آن ها عبارتند از: نما، میانه، میانگین، و منظور از شاخص های پراکندگی آن است که شاخص مورد نظر جایگاه خاصی در توزیع فراوانی ندارد. به عبارت دیگر شاخصی مورد بررسی می تواند در ابتدا، وسط، و یا انتهای توزیع یافت شود که معروف ترین آن ها عبارتند از: دامنه تغییرات R، انحراف چارکی $\frac{Q_3 - Q_1}{2}$، انحراف متوسط $\frac{|X - \bar{X}|}{n}$، واریانس S^2، انحراف معیار S

☑ شاخص های مرکزی نسبت به ۴ عمل اصلی ریاضی حساس بوده به طوری که با کمترین میزان تغییر در توزیع فراوانی آن شاخص نیز تغییر می کند.

☒ شاخص های پراکندگی مخصوصا واریانس و انحراف معیار نسبت به ۴ عمل اصلی ریاضی حساس نمی باشند

محاسبه‌ی نما بدون استفاده از جدول توزیع فراوانی

برای محاسبه نما با یک شیوه کافی است به ظاهر اعداد توجه نماییم و عددی که بیش از سایر اعداد تکرار داشته باشد را به عنوان مد معرفی می نماییم مانند

۱۲، ۱۲، ۱۶، ۱۲، ۷، ۱۳، ۱۲، ۱۴، ۱۵، ۸ → توزیع فراوانی

$M_o = 12$ → مد

۶، ۲، ۶، ۷، ۳، ۵، ۶، ۶، ۷ → $M_o = 6$

گاه توزیع می تواند بی شمار مد داشته باشد مانند

۲، ۷، ۳، ۵، ۶، ۶، ۷ → $M_o = 6, 7$

گاه توزیع می تواند مد نداشته باشد مانند

۱، ۲، ۳، ۴، ۵، ۶ → $M_o = \emptyset$

محاسبه میانه بدون استفاده از جدول توزیع فراوانی

ابتدا داده‌ها را به صورت صعودی یا نزولی مرتب می کنیم

Scanned with CamScanner

Date Subject

برای محاسبه ی این شاخص ابتدا کلیه اعداد را مرتب کرده و سپس به روش های زیر اقدام می نماییم

الف) اگر تعداد اعداد فرد باشد پس از مرتب نمودن شروع به حذف نمودن اعداد از دو سمت توزیع می نماییم و در نهایت عددی که باقی ماند به عنوان میانه معرفی می شود مانند

۶-۷-۹-۱۲-۱۴-۶-۱۵-۱۰-۱۶

۶-۶-۷-۹-(۱۰)-۱۴-۱۴-۱۵-۱۶ ⟶ میانه = ۱۰

ب) هرگاه تعداد اعداد زوج بوده در این صورت پس از مرتب نمودن اعداد شروع به حذف نمودن اعداد ۲ سمت توزیع می نماییم و در نهایت ۲ عدد باقی مانده را با هم جمع و تقسیم بر ۲ می کنیم و عدد بدست آمده به عنوان میانه معرفی می شود. مانند

۵-۶-۱۰-۱۲-۱۴-۹ ⟶ ۵-۶-۹-۱۰-۱۴-۱۴

$$\frac{9+10}{2} = \text{میانه } 9/5$$

ج) درصورتی که تعداد اعداد زوج بوده و اعداد دارای تکرارهای یکسان نباشد به شیوه زیر عمل می نماییم.

۶-۷-۸-۸-۸-۱۰-۱۲-۱۴-۱۴-۱۵

$$\text{عدد نوجه میانی} = \text{عدد میانی} + \frac{\text{تعداد تکرار عدد میانی} \times \text{اختلاف ۲ عدد میانی}}{\text{مجموع تکرار اعداد میانی}}$$

$$\begin{cases} 10 - \frac{2\times 2}{5} = 10 - 0/8 = 9/2 \\ \text{یا} \\ 8 + \frac{3\times 2}{5} = 8 + 1/2 = 9/2 \end{cases}$$

مثال بعدی)

۸-۹-۹-۹-۹-۱۳-۱۳-۱۴-۱۵-۱۶

$$9 + \frac{4\times 4}{7} = 11/28$$

$$13 - \frac{3\times 4}{7} = 11/29$$

Kian SNAD

Scanned with CamScanner

Date: Subject:

محاسبه میانگین بدون استفاده از جدول توزیع فراوانی

برای این منظور کافی است تمامی اعداد را با یکدیگر جمع سپس تقسیم بر تعداد نماییم به عبارت دیگر از فرمول روبه رو استفاده می کنیم

$$\bar{X} = \frac{\sum X}{N}$$

محاسبه نما یا مد با استفاده از جدول توزیع فراوانی

مد = بیشترین تکرار - برای محاسبه نمایه شاخص های آماری با استفاده از جدول می بایست از فرمول زیر استفاده نماییم. برای محاسبه نما یا مد از فرمول زیر استفاده می کنیم

$$Mo = Li + \frac{d_1}{d_1 + d_2} \times i$$

- Li: کران پایین طبقه ای که مد در آن قرار دارد.
- d_1: اختلاف فراوانی طبقه ای که مد در آن قرار دارد با طبقه ماقبلش
- d_2: اختلاف فراوانی طبقه ای که مد در آن قرار دارد با طبقه مابعدش
- i: فاصله طبقات

محاسبه میانه با استفاده از جدول توزیع فراوانی

برای محاسبه میانه ابتدا ستونی تحت عنوان cf در جدول اضافه می شود برای تکمیل این ستون کافی است اولین عدد موجود در ستون f در اولین ردیف ستون cf قرار گیرد در ادامه آن عدد با دومین عدد موجود در ستون f جمع شده و نتیجه بدست آمده در دومین ردیف cf قرار می گیرد و این عمل را برای تمامی طبقات انجام می دهیم سپس $\frac{N}{2}$ عددها را محاسبه می نماییم در ادامه جایگاه عدد حاصل از $\frac{N}{2}$ را در ستون cf می یابیم تا طبقه ملاک برای استفاده از فرمول مشخص گردد و در نهایت از فرمول زیر استفاده می نماییم

$$Md = Li + \frac{\frac{N}{2} - cf_b}{f} \times i$$

- Li: کران پایین طبقه ای که میانه در آن قرار دارد.
- f: اونیکه میانه در آن قرار دارد.
- cf_b: فراوانی ماقبل

Scanned with CamScanner

برای محاسبه چارک ها همانند محاسبه میانه از فرمول استفاده نمود و تنها تفاوت مورد نیاز تفسیری نماید

به طوری که در چارک اول به دنبال نقطه $\frac{N}{4}$ بوده یعنی فرمول به شیوه زیر تفسیری نماید.

$$Q_1 = L_1 + \frac{\frac{N}{4} - cf_b}{f_0} \times i$$

← فراوانی ماقبل

و در چارک سوم ($\frac{3N}{4}$) نقطه به $\frac{3N}{4}$ تفسیر می‌شود و فرمول به شیوه زیر می‌شود

$$Q_3 = L_1 + \frac{\frac{3N}{4} - cf_b}{f_0} \times i$$

پس از محاسبه‌ی چارک اول و سوم می‌توانیم با استفاده از فرمول زیر انحراف چارکی توزیع را محاسبه نماییم

$$Q = \frac{Q_3 - Q_1}{2}$$

↓ نیوخالی نماد انحراف چارکی است

محاسبه میانگین با استفاده از جدول توزیع فراوانی

برای محاسبه میانگین کافی است ۲ ستون تحت عنوان X و fx به جدول اضافه می‌نماییم برای تکمیل ستون X نماینده واقعی اعداد (طبقه) محسوب می‌شود، کافی است آن ۲ عدد قابل مشاهده در ستون L_1 را با یکدیگر جمع سپس تقسیم بر ۲ می‌کنیم هم‌چنین برای تکمیل ستون fx کافی است اعداد ستون f را در اعداد ستون X ضرب نماییم و در ادامه اعداد ستون fx را با یکدیگر جمع می‌نماییم تا $\sum fx$ حاصل شود سپس از فرمول زیر استفاده می‌کنیم

$$\bar{X} = \frac{\sum f.x}{N}$$

محاسبه واریانس و انحراف معیار با استفاده از جدول توزیع فراوانی

برای محاسبه این ۲ شاخص ابتدا باید میانگین اعداد را مقایسه کنیم در صورت آن‌که میانگین به دست آمده اعشاری بود آن را رند می‌نماییم سپس ۳ ستون تحت عنوان x، x^2، fx^2 به جدول اضافه می‌نماییم. برای تکمیل ستون x از قاعده‌ی روبه رو استفاده می‌نماییم $(X - \bar{X})$

برای تکمیل ستون x^2 کافی است اعداد ستون x را به توان ۲ برسانیم هم‌چنین برای تکمیل ستون fx^2 کافی است اعداد ستون f را در اعداد ستون x^2 ضرب نموده و در نهایت اعداد آن ستون را باید یکدیگر جمع می‌نماییم تا $\sum fx^2$ حاصل شود. در نهایت از فرمول زیر برای محاسبه واریانس

Scanned with CamScanner

$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N-1}$ و از فرمول $S = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N-1}}$ برای محاسبه انحراف معیار استفاده می‌کنیم

نکته: قاعده رند کردن اعداد.

برای رند کردن اعداد (گرد کردن اعداد) به اولین عدد بعد از اعشار توجه می‌نماییم و آن عدد را با ۵ عدد مقایسه قرار می‌دهیم. اگر عدد موق از ۵/۰ بیشتر بود رو به بالا رند می‌نماییم یعنی اعشار را حذف نموده و یک واحد به عدد صحیح قبل از اعشار اضافه می‌کنیم مانند

$۸٫۶ \simeq ۹$

اگر اولین عدد بعد اعشار کوچکتر از ۵/۰ بود اعشار را حذف نموده و عدد صحیح باقی‌مانده را انتخاب می‌نماییم مانند

$۷٫۴ \simeq ۷$ $\quad$ $۲۸٫۵۰۰۰۰۲ \simeq ۲۹$

اگر اولین عدد بعد از اعشار خود ۵/۰ بوده و بعد از آن عددی وجود نداشته باشد در این صورت به اولین عدد قبل از اعشار توجه می‌کنیم اگر آن عدد فرد بود رو به بالا رند می‌کنیم مانند $۲۱٫۵ \simeq ۲۲$ و اگر آن عدد زوج بود رو به پایین رند می‌نماییم مانند $۱۸٫۵ \simeq ۱۸$

محاسبه واریانس و انحراف معیار بدون استفاده از جدول توزیع فراوانی

ابتدا تمامی اعداد را با هم جمع نموده تا $\sum x$ حاصل شود سپس تمامی اعداد را به توان ۲ می‌رسانیم و باید نیز جمع می‌کنیم تا $\sum x^2$ حاصل شود و در نهایت از فرمول زیر استفاده می‌نماییم

$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N-1} \qquad S = \sqrt{S^2}$$

مانند

۳-۵-۷-۹-۱۱-۱۲

x	x^2
۳	۹
۵	۲۵
۷	۴۹
۹	۸۱
۱۱	۱۲۱
۱۲	۱۴۴
۴۷	۴۲۹

$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N-1} = \frac{۴۲۹ - \frac{(۴۷)^2}{۶}}{۶-۱} \qquad S = \sqrt{S^2} = \sqrt{۱۲٫۱۶}$$

$$S = ۳٫۴۸$$

$$S^2 = \frac{۴۲۹ - \frac{۲۲۰۹}{۶}}{۵} = \frac{۴۲۹ - ۳۶۸٫۱۵}{۵} = \frac{۶۰٫۸۴}{۵} = ۱۲٫۱۶$$

PAPCO

Scanned with CamScanner

Subject:
Date

۵ - ۷ - ۹ - ۱۱ - ۱۳ - ۱۴

X	X^2
۵	۲۵
۷	۴۹
۹	۸۱
۱۱	۱۲۱
۱۳	۱۴۹
۱۴	۱۹۶
۵۹	۴۸۱

$$S^2 = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}$$

$$S^2 = \frac{۴۸۱ - \frac{(۵۹)^2}{۶}}{۶-1} = \frac{۴۸۱ - \frac{۳۴۸۱}{۶}}{۵}$$

$$S^2 = \frac{۴۸۱ - ۵۸۰/۱۶}{۵} = ۱۲/۱۶$$

$$S = \sqrt{S^2} = \sqrt{۱۲/۱۶} = ۳/۴۸$$

$$S = ۳/۴۸$$

کلیه شاخص های پراکندگی

مثال زیر را حل کنید؟

cl	f	cf	X	f.X	$x-\bar{x}$	x^2	$f.x^2$
۵-۹	۳	۳	۷	۲۱	۷-۲۰=-۱۳	۱۶۹	۵۰۷
۱۰-۱۴	۸	۱۱	۱۲	۹۶	-۸	۶۴	۵۱۲
۱۵-۱۹	۴	۱۵	۱۷	۶۸	-۳	۹	۳۶
۲۰-۲۴	۶	۲۱	۲۲	۱۳۲	۲	۴	۲۴
۲۵-۲۹	۸	۲۹	۲۷	۲۱۶	۷	۴۹	۳۹۲
۳۰-۳۴	۳	۳۲	۳۲	۹۶	۱۲	۱۴۴	۴۳۲
	۳۲			۶۲۹			۱۹۰۳

میانه: $\frac{۵+۹}{۲} = ۷$ → وسط جدول برای ترتیب x → $\frac{۱۰+۱۴}{۲} = ۱۲$

۹/۵ ← کران پایین ۱۰ ← کران بالا ۱۰/۵

بدست آوردن مد: $Mo = Li + \frac{\Delta 1}{\Delta 1 + \Delta 2} \times i$

$Mo = ۹/۵ + \frac{۵}{۵+۴} \times ۵$

$Mo = ۱۲/۲۷$

Scanned with CamScanner

$$Mo = ۲۴/۵ + \frac{۲}{۲+۵} \times ۵$$

$$Mo = ۲۴/۵ + ۰/۲۸۵$$

$$Mo = ۲۵/۹۲$$

برای بدست آوردن میانه $Mo = Li + \frac{\frac{N}{۲} - cf_b}{f} \times i$

$\frac{N}{۲} = \frac{۳۲}{۲} = ۱۶ \longrightarrow Mo = ۱۹/۵ + \frac{۱۶ - ۱۵}{۶} \times ۵$

$Mo = ۲۰/۳۴$

برای محاسبه چارک $Q_۲ = Li + \frac{\frac{N}{۴} - cf_b}{f} \times i$

$\frac{N}{۴} = \frac{۳۲}{۴} = ۸ \longrightarrow Q_۲ = ۹/۵ + \frac{۸ - ۳۰}{۸} \times ۵$

$Q_۲ = ۹/۵ + (۰/۶۲۵ \times ۵) = ۱۲/۶۲۵$

$۳/۱۲۵$

برای بدست آوردن چارک سوم $\longrightarrow Q_۳ = Li + \frac{\frac{۳N}{۴} - cf_b}{f} \times i$

$\frac{۳N}{۴} = \frac{۳ \times ۳۲}{۴} = ۲۴ \qquad Q_۳ = ۲۴/۵ + \frac{۲۴ - ۲۱}{۸} \times ۵$

$Q_۳ = ۲۴/۵ + ۱/۸۷ = ۲۶/۳۷$

انحراف چارکی $\longrightarrow Q = \frac{Q_۳ - Q_۱}{۲} = \frac{۲۶/۳۷ - ۱۲/۶۲۵}{۲} = ۶/۸۷$

میانگین $\longrightarrow \bar{X} = \frac{\Sigma fx}{N} = \frac{۶۲۹}{۳۲} = ۱۹/۶۵$

واریانس $\longrightarrow S^۲ = \frac{\Sigma f.x^۲}{N} = \frac{۱۹۰۳}{۳۲} = ۵۹/۴۶$

انحراف معیار $\longrightarrow S = \sqrt{\frac{\Sigma f.x^۲}{N}} = \sqrt{۵۹/۴۶} = ۷/۷۱۱$

PAPCO

Scanned with CamScanner

مثال ③

cf	f	f_x	x	$x^*-\bar{x}$	$f_x x^2$	cf	x^2
۶_۱۲	۴	۳۶	۹	-۱۴	۷۸۴	۴	۱۹۶
۱۳_۱۹	۶	۹۶	۱۶	-۷	۲۹۴	۱۰	۴۹
۲۰_۲۶	۱۰	۲۳۰	۲۳	۰	۰	۲۰	۰
۲۷_۳۳	۷	۲۱۰	۳۰	۷	۳۴۳	۲۷	۴۹
۳۴_۴۰	۳	۱۱۱	۳۷	۱۴	۵۸۸	۳۰	۱۹۶
Σ	۳۰	۶۸۳			۲۰۰۹		

$$\bar{X} = \frac{\Sigma f.x}{N} = \frac{۶۸۳}{۳۰} = ۲۲,۷۶ \simeq ۲۳$$

$$S^2 = \frac{\Sigma f.x^2}{N} = \frac{۲۰۰۹}{۳۰} = ۶۶,۹۶$$

$$S = \sqrt{\frac{\Sigma f.x^2}{N}} = \sqrt{۶۶,۹۶} = ۸,۱۸$$

$$Mo = Li + \frac{d_1}{d_1+d_2} \times i = ۱۹,۵ + \frac{۴}{۴+۳} \times ۷ = ۲۳,۵$$

$$\boxed{Mo = ۲۳,۵}$$

$$Md = Li + \frac{\frac{N}{۲} - cf_b}{f} \times i = ۱۹,۵ + \frac{۱۵-۱۰}{۱۰} \times ۷ = ۲۳$$

$$\boxed{Md = ۲۳} \qquad \frac{N}{۲} = \frac{۳۰}{۲} = ۱۵$$

هرگاه بخواهیم به بررسی ارتباط موجود بین ۲ یا چند دسته از اطلاعات بپردازیم و میزان و جهت رابطه را تعیین کنیم می توانیم از ضرایب هم بستگی استفاده نماییم. ضریب هم بستگی را همواره با r نشان داده و همیشه عددی است بین مثبت و منفی ۱ ($-۱ \leq r \leq +۱$) هم بستگی مثبت ۱ اشاره به وجود یک رابطه ی کامل مستقیم دارد. یعنی اگر x افزایش یابد y نیز افزایش می یابد مانند رابطه بین میزان دارایی شرکت و سود آن. هم بستگی منفی ۱ به وجود یک رابطه کامل معکوس اشاره می کند.

Scanned with CamScanner

یعنی اگر x افزایش یابد y کاهش می یابد مانند رابطه بین عرضه و تقاضا. هم بستگی می تواند صفر نیز باشد که به آن عدم هم بستگی می گویند و همواره هم بستگی نیز می تواند ابعادی بین مثبت و منفی ۱ باشد که به این گونه از هم بستگی ها، هم بستگی ناقص می گویند.

ناقص معکوس
(-۱) ——— -۱/۳۲ ——— ۰ ——— ۰/۸۱ ——— (+۱)
کامل معکوس ↓ عدم هم بستگی ناقص مستقیم کامل مستقیم

برای محاسبه هم بستگی می توان از ضرایب مختلف هم بستگی استفاده کرد - معروف ترین آن ها عبارتند از پیرسون - اسپیرمن - کندال

الف) ضریب هم بستگی پیرسون: هرگاه بخواهیم به بررسی رابطه ی بین ۲ نفر یا چند دسته از داده های کمی بپردازیم از این ضریب می توان استفاده کرد. به عبارتی از فرمول زیر استفاده می کنیم

ضریب هم بستگی پیرسون (محاسبه با استفاده از داده های خام)

$$r_{xy} = \frac{N.\Sigma xy - (\Sigma x).(\Sigma y)}{\sqrt{[N.\Sigma x^2 - (\Sigma x)^2][N.\Sigma y^2 - (\Sigma y)^2]}}$$

برای تفسیر ضریب هم بستگی پیرسون ابتدا درجه آزادی اعداد را با استفاده از $df = N - 1$ بدست می آوریم سپس براساس سطح اطمینان ارائه شده و درجه آزادی بدست آمده از جداول انتهای کتاب های آمار اقدام به یافتن عدد ملاک جهت تفسیر می نماییم هرگاه ضریب هم بستگی محاسبه شده بزرگتر از ضریب ملاک در جدول بود می توان نتیجه گرفت که بین ۲ دسته از اطلاعات رابطه معناداری وجود دارد

در آمار ۲ مدل اطمینان داریم:

df: درجه آزادی ← ① محاسبه ← ۰/۹۵ ← خطا ← ۰/۰۵

② تفسیر ← ملاک ← ۰/۹۹ ← خطا ← ۰/۰۱

$df = N - 1$

↓

درجه آزادی ضریب هم بستگی

Scanned with CamScanner

مثال زیر را حل کنید

X	Y	X^2	Y^2	XY
۴	۱۰	۱۶	۱۰۰	۴۰
۶	۱۵	۳۶	۲۲۵	۹۰
۲	۸	۴	۶۴	۱۶
۷	۱۶	۴۹	۲۵۶	۱۱۲
۱۰	۱۹	۱۰۰	۳۶۱	۱۹۰
۱۴	۲۰	۱۹۶	۴۰۰	۲۸۰
۴۳	۸۸	۴۰۱	۱۴۰۶	۷۲۸

$\Sigma X = ۴۳$

$\Sigma Y = ۸۸$

$\Sigma X^2 = ۴۰۱$

$\Sigma Y^2 = ۱۴۰۶$

$\Sigma XY = ۷۲۸$

$$r_{xy} = \frac{N.\Sigma XY - (\Sigma X).(\Sigma Y)}{\sqrt{[N\Sigma X^2 - (\Sigma X)^2][N.(\Sigma Y)^2 - (\Sigma Y)^2}}$$

$$r_{xy} = \frac{(۶ \times ۷۲۸) - (۴۳)(۸۸)}{\sqrt{[(۶ \times ۴۰۱) - (۴۳)^2][(۶ \times ۱۴۰۶) - (۸۸)^2}} =$$

$$r_{xy} = \frac{۴۳۶۸ - ۳۷۸۴}{\sqrt{[۲۴۰۶ - ۱۸۴۹][۸۴۳۶ - ۷۷۴۴]}} = \frac{۵۸۴}{\sqrt{۳۱۵۴۴۴}}$$

$$r_{xy} = \frac{۵۸۴}{۶۲۰۱۸۴۱} \longrightarrow \boxed{r_{xy} = ۰/۹۴}$$

ناقص مستقیم

☑ اگر صورت کسر منفی باشد هم بستگی منفی می شود

محقق قصد دارد به بررسی رابطه بین سود و دارایی شرکت بپردازد. برای این منظور به جمع آوری داده های مورد نیاز می پردازد و همت شرکت را به عنوان نمونه خود انتخاب می کند پس از بررسی صورت گرفته داده های جدول زیر حاصل می شود مطلوب است با اطمینان ۹۹ درصد به بررسی رابطه بین سود و دارایی شرکت بپردازید.

$۰/۷۳ = ۰/۹۹$ جدول

Scanned with CamScanner

سود	درآمد	X^2	Y^2	XY
۵	۱۵	۲۵	۲۲۵	۷۵
۷	۱۰	۴۹	۱۰۰	۷۰
۱۲	۲۰	۱۴۴	۴۰۰	۲۴۰
۱۹	۴۵	۳۶۱	۲۰۲۵	۸۵۵
۱۴	۳۰	۱۹۶	۹۰۰	۴۲۰
۲۰	۵۰	۴۰۰	۲۵۰۰	۱۰۰۰
۱۷	۴۰	۲۸۹	۱۶۰۰	۶۸۰
۹۴	۲۱۰	۱۴۶۴	۷۷۵۰	۳۳۴۰

$$r_{xy}=\frac{N.\Sigma XY-(\Sigma X)(\Sigma Y)}{\sqrt{[N.(\Sigma X^2)-(\Sigma X)^2][N.\Sigma Y^2-(\Sigma Y)^2}}$$

$$r_{xy}=\frac{(7\times 3340)-(94\times 210)}{\sqrt{[(7\times 1464)-(94)^2][(7\times 7750)-(210)^2}}$$

$$r_{xy}=\frac{23380-19740}{\sqrt{(10248-8836)(54250-44100)}}$$

$$r_{xy}=\frac{3640}{\sqrt{(1412)(10150)}}$$

$$r_{xy}=\frac{3640}{\sqrt{14323180}}$$

$$r_{xy}=\frac{3640}{3785/73}\doteq 0/96$$

$r_{xy}=0/96$

باتوجه به آن که میزان ضریب هم بستگی محاسبه شده ($r=0/96$) بزرگتر از مقدار جدول ($r=0/873$) می باشد بنابراین می توان با اطمینان ۹۹٪ بیان داشت بین سود و درآمد شرکت یک رابطه معنادار مستقیم ناقص وجوددارد.

فرضیه و انواع آن

فرضیه یک بیانیه خبری و یک حدس عالمانه در بین ۲ یا چند پدیده می باشد که اشاره به وجود یا عدم وجود رابطه، تفاوت و تاثیر در بین متغیرها دارد و همواره در انتهای آن نیز نقطه قرار می گیرد و به صورت یک جمله مثبت بیان می شود. فرضیه ای که به صورت مثبت بیان شده فرض تحقیق یا H_1 نام دارد و در مقابل آن نیز فرض خلاف یا صفر یا پوچ یا آماری قرار دارد که دلالت بر عدم وجود رابطه، تفاوت تاثیر در بین متغیرها دارد و با H_0 نشان می دهند. فرض تحقیق از سوال تحقیق ایجاد می شود به طوری که سوال مورد بررسی از حالت سوالی خارج شده و به صورت یک جمله خبری بیان می شود.

Scanned with CamScanner

مثال) آیا بین سبک تدریس علم با یادگیری دانش آموزان رابطه وجود دارد؟

فرض تحقیق

فرضیه

H_1 بین سبک تدریس علم با یادگیری دانش آموزان رابطه وجود دارد

H_0 بین سبک تدریس علم با یادگیری دانش آموزان رابطه وجود ندارد

مثال صفحه قبل ← آیا بین سود و میزان دارایی شرکت رابطه وجود دارد؟

H_1 = بین سود و میزان دارایی شرکت رابطه وجود دارد ۲= ۰/۹۶

H_0 = بین سود و میزان دارایی شرکت رابطه وجود ندارد ۲= ۰/۷۳

ب) ضریب هم بستگی اسپیرمن: اگر بخواهیم به بررسی رابطه بین ۲ دسته از اطلاعات که دارای مقیاس رتبه ای باشند بپردازیم (بررسی روابط در بین کدها) می توان از ضریب هم بستگی اسپیرمن استفاده کرد و برای این منظور از فرمول زیر استفاده می نماییم:

$$r_{hs} \text{ یا } r_s = \frac{1 - 6\Sigma D^2}{N(N^2-1)}$$

پژوهشگری قصد دارد به بررسی رابطه بین سود شرکت و میزان دارایی آن بپردازد. برای این منظور اقدام به جمع آوری داده های مورد نیاز در بین شرکت های مورد بررسی می پردازد. پس از جمع آوری داده های مورد نیاز نتایج جدول زیر بدست می آید مطلوب است با اطمینان ۹۹٪ به بررسی رابطه رتبه ای بین ۲ دسته از اطلاعات بپردازید. (۲= ۰/۴۱)

آیا بین سود شرکت و میزان دارایی آن رابطه وجود دارد؟

✓ H_1 = بین سود شرکت و میزان دارایی آن رابطه وجود دارد

✗ H_0 = بین سود شرکت و میزان دارایی آن رابطه وجود ندارد

Scanned with CamScanner

X	Y	R_x	R_y	D	D^2
۱۴	۱۸	۵	۵	۰	۰
۱۹	۴۲	۳	۲	۱	۱
۷	۲۱	۶	۴	۲	۴
۳	۹	۷	۶	۱	۱
۱۷	۴	۴	۷	-۳	۹
۲۵	۳۵	۱	۳	-۲	۴
۲۲	۴۸	۲	۱	۱	۱
					۲۰

r_{ho} یا $r_b = 1 - \frac{۶\Sigma D^2}{N(N^2-1)}$

r_{ho} یا $r_b = 1 - \frac{۶(۲۰)}{۷(۷^2-1)}$

r_{ho} یا $r_b = 1 - \frac{۱۲۰}{۷\times ۴۸}$

r_{ho} یا $r_b = 1 - \frac{۱۲۰}{۳۳۶} = 1 - ۰/۳۵$

r_{ho} یا $r_b = ۰/۶۵$

تفسیر $r_{ho} = ۰/۴۱ < r_{ho} = ۰/۶۵$

رابطه معناداری وجود دارد یعنی فرض تحقیق تایید و فرض خلاف رد می شود یعنی با افزایش میزان دارایی شرکت سود آن نیز افزایش می یابد

اگر r_{ho} کمتر از ۴۱ درصد می شد فرضیه Ho وجود ندارد جواب ما می شد

پژوهشگری قصد دارد به بررسی رابطه بین هزینه خانوار و امید به زندگی بپردازد. برای این منظور اقدام به جمع آوری داده های مورد نیاز در بین شرکت های مورد بررسی می پردازد پس از جمع آوری داده ها نتایج جدول زیر بدست می آید مطلوب است با اطمینان ۰/۴۸=r۰ بررسی رابطه رتبه ای بین ۲ دسته از اطلاعات بپردازید.

✓ بین هزینه خانوار و امید به زندگی رابطه وجود دارد = H_1

✗ بین هزینه خانوار و امید به زندگی رابطه وجود ندارد = H_0

r_{ho} یا $r_b = 1 - \frac{۶\Sigma D^2}{N(N^2-1)}$

r_{ho} یا $r_b = 1 - \frac{۶\times ۱۷/۵}{۸((۸)^2-1)} = 1 - \frac{۱۰۵}{۸\times ۶۳}$

P4PCO

Scanned with CamScanner

$r_{hO} = ۱ - \frac{۱۰۵}{۵۰۴}$

$r_{hO} = ۱ - ۰/۲ \longrightarrow r_{hO} = ۰/۸$

$r_{hO} = ۰/۸ > r_{hO} = ۰/۴۸$
(مقدار r) (جدول)

رابطه وجود دارد و ۰ الحاق رد می‌شود
و ۱ قبول می‌شود و مرحله تحقیق پایه و مرحله
حذف رد می‌شود

هزینه ماهیانه X	رتبه‌ی Y	R_X	R_Y	D	D^2
۱۰	۱۵	۶/۵	۵/۵	۱	۱
۲۵	۱۹	۲/۵	۳/۵	-۱	۱
۶۰	۲۰	۱	۱	۰	۰
۲۵	۲۰	۲/۵	۲	۰/۵	۰/۲۵
۱۷	۱۵	۴	۵/۵	-۱/۵	۲/۲۵
۴	۱	۸	۸	۰	۰
۱۳	۸	۵	۷	-۲	۴
۱۰	۱۹	۶/۵	۳/۵	۳	۹
					۱۷/۵

رتبه ۳ — رتبه ۳ — $\frac{۳+۲}{۲}$
$\frac{۶+۷}{۲}$

آزمون های مقایسه‌ای: هرگاه بخواهیم به بررسی تفاوت میانگین یک متغیر با یک جامعه یا یک یا چند گروه بپردازیم می‌توان از این آزمون ها استفاده کرد. این آزمون ها در زمره‌ی آزمون های پارامتریک قرار گرفته بنابراین یکی از شروط استفاده از آن ها نرمال بودن متغیر وابسته می‌باشد معروف‌ترین این خانواده از آزمون ها عبارتند از: آزمون t تک گروهی - t مستقل - t همتا - F تحلیل واریانس

الف) آزمون t تک گروهی: هرگاه بخواهیم به مقایسه میانگین یک گروه یا یک دسته از اطلاعات با میانگین جامعه بپردازیم می‌توان از این آزمون استفاده کرد برای استفاده از این آزمون از مدل زیر استفاده می‌نماییم

$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}}$

$\bar{X}$ → میانگین اعداد
μ_0 → میانگین جامعه
$S_{\bar{X}}$ → خطای انحراف میانگین

انحراف معیار $S_{\bar{X}} = \frac{S}{\sqrt{n}}$
S → انحراف معیار
n → تعداد نمونه

Scanned with CamScanner

Subject:
Date:

۲۰

مدیر یک کارگاه تولیدی ادعا می‌کند میانگین حقوق روزانه پرداختی به ۹ واحد تحت سرپرستی حد ۲۰۰ هزار تومان است پژوهشگری قصد دارد به بررسی این ادعا بپردازد بنابراین ۹ نفر از کارگران را به صورت تصادفی جهت بررسی این ادعا انتخاب می‌کند پس از جمع آوری داده های مورد نیاز مشخص شد که میانگین و انحراف معیار حقوق پرداخت شده به ترتیب برابر با ۱۸۰ و ۱۴ می‌باشد فرض است با اطمینان ۹۵ درصد به بررسی این ادعا که آیا بین میانگین حقوق ارائه شده و میانگین جامعه تفاوت وجود دارد میپردازیم

$t_{ملاک} = ۲,۷۲$

$n = ۹$

$S = ۱۴$

$\bar{X} = ۱۸۰$

$\mu_0 = ۲۰۰$

$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}}$

$S_{\bar{X}} = \frac{S}{\sqrt{n}}$

$S_{\bar{X}} = \frac{۱۴}{\sqrt{۹}} = ۴,۶۷$

$t = \frac{۱۸۰ - ۲۰۰}{۴,۶۷} = \frac{-۲۰}{۴,۶۷}$

$t = -۴,۲۸ > t_{ملاک} = ۲,۷۲$

منفی نشان دهنده اختلاف و تغییر

✓ H_1 = بین میانگین حقوق ارائه شده با میانگین جامعه تفاوتی وجود دارد

X H_0 = بین میانگین حقوق ارائه شده با میانگین جامعه تفاوتی وجود ندارد

تفسیر: با توجه به آن که t محاسبه شده از t جدول بزرگتر می باشد $t = -۴,۲۸ > t = ۲,۷۲$ می‌توان نتیجه گرفت که بین میانگین محاسبه شده و میانگین جدول از لحاظ آماری تفاوت معناداری مشاهده می‌شود بنابراین فرض تحقیق تایید و فرض خلاف رد می‌شود از طرفی با توجه به منفی بودن میزان t می توان نتیجه گرفت که وضع موجود ارائه حقوق و مزایا به کارگران با وضع مطلوب تفاوتی دارد و کارگران از حقوق مطلوبی برخوردار نمی باشند

پژوهشگری قصد دارد به بررسی این ادعا بپردازد که میانگین عملکرد تحصیلی دانش آموزان یک مدرسه برابر با ۱۵ می باشد برای این منظور وی ۲۹ نفر از دانش آموزان را به صورت تصادفی انتخاب و عملکرد تحصیلی آنان را محاسبه می نماید. پس از بررسی صورت گرفته مشخص شد که میانگین و انحراف معیار عملکرد تحصیلی دانش آموزان برابر است با ۱۴ و ۱۰ مطلوب است با اطمینان ۹۹٪ به بررسی این ادعا بپردازید $t_{ملاک} = ۲,۹۶$

X H_1 = بین میانگین عملکرد تحصیلی دانش آموزان با میانگین جامعه تفاوتی وجود دارد

✓ H_0 = بین میانگین عملکرد تحصیلی دانش آموزان با میانگین جامعه تفاوتی وجود ندارد

Scanned with CamScanner

Subject:
Date

$\bar{X} = ۱۴$ $n = ۴۹$

$\mu_0 = ۱۵$ $S = ۱۰$

$$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}} = \frac{۱۴ - ۱۵}{۱,۴۲} = \frac{-۱}{۱,۴۲}$$

چون t جدول از t محاسبه شده کمتر است می‌توان نتیجه گرفت
بین میانگین نمونه و میانگین جامعه از لحاظ آماری تفاوت
وجود دارد. بنابراین فرض خلاف تایید و فرض تحقیق رد می‌شود

$t = -۰/۷۰ < t = ۲,۶۶$

$$S_{\bar{X}} = \frac{۱۰}{\sqrt{۴۹}} = ۱,۴۲$$

ب) آزمون t گروه های مستقل: هرگاه بخواهیم به مقایسه ی میانگین در ۲ گروه که مستقل از یکدیگر باشند بپردازیم می‌توان از این آزمون استفاده کرد. این آزمون یک آزمون پارامتریک محسوب شده و می‌بایست متغیرهای مورد بررسی حتما از همگونه‌ی نرمال بودن برخوردار باشند. برای استفاده از این آزمون از فرمول زیر استفاده می‌کنیم

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

← واریانس گروه اول

↓ تعداد دسته گروه اول

پژوهشگری قصد دارد به میزان سود ۱۰ شرکت دولتی و ۱۰ شرکت غیر دولتی بپردازد برای این منظور پس از بررسی صورت های مالی شرکت های مورد بررسی مشخص شد که میانگین و واریانس شرکت های دولتی به ترتیب برابر است با ۱۲۰۰ و ۶۵ و از طرفی میانگین و واریانس سود شرکت های غیردولتی به ترتیب برابر است با ۱۸۰۰ و ۵۵ است مطلوب است با اطمینان ۹۵ره درصد به مقایسه سود شرکت های دولتی و غیردولتی بپردازید

H_1 = بین سود شرکت های دولتی و غیردولتی تفاوت وجود دارد

H_0 = بین سود شرکت های دولتی و غیردولتی تفاوت وجود ندارد

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}} = \frac{۱۲۰۰ - ۱۸۰۰}{\sqrt{\frac{۶۵}{۱۰} + \frac{۵۵}{۱۰}}} = \frac{-۶۰۰}{\sqrt{\frac{۱۲۰}{۱۰}}} = \frac{-۶۰۰}{\sqrt{۱۲}}$$

$$t = \frac{-۶۰۰}{۳,۴۶} = -۱۷۳,۴۱$$

PAPCO

Scanned with CamScanner

$t_0 = -۱۷۲,۴۱ > t = ۱,۹۶$

چون t محاسبه شده از t جدول بزرگتر است پس تفاوت وجود دارد پس فرض تحقیق تایید و فرض خلاف رد می شود و از طرفی با توجه به منفی بودن میزان t می توان نتیجه گرفت وضع مطلوب سود شرکت های دولتی و غیردولتی با وضع مطلوب تفاوتی دارد و سود های شرکت های دولتی و غیردولتی از وضع مطلوبی برخوردار نیست

غیردولتی (X_2)	دولتی (X_1)
$\bar{X} = ۱۸۰۰$	$\bar{X} = ۱۲۰۰$
$s_2 = ۵۵$	$s_1^2 = ۲۵$
$N = ۱۰$	$N = ۱۰$

محققی قصد دارد به مقایسه شغلی کارکنان یک شرکت بپردازد بنابراین ۱۲ کارمند زن و ۱۲ کارمند مرد را برای انجام این تحقیق انتخاب می کند در ادامه با توضیح یک پرسشنامه رضایت شغلی در بین آن ها میزان رضایت شغلی شان را ارزیابی می کند پس از تحلیل داده ها مشخص شد که میانگین و واریانس رضایت شغلی زنان به ترتیب برابر با ۶۲ و ۱۰ و میانگین و واریانس رضایت شغلی مردان به ترتیب برابر با ۵۰ و ۱۱ است مطلوب است مقایسه رضایت شغلی زنان و مردان با یکدیگر بپردازید. t ملاک = ۱,۹۶

مردان	زنان
$n_2 = ۱۲$	$n_1 = ۱۲$
$\bar{X}_2 = ۵۰$	$\bar{X}_1 = ۶۲$
$S_2^2 = ۱۱$	$S_1^2 = ۱۰$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}} = \frac{۶۲ - ۵۰}{\sqrt{\frac{۱۰}{۱۲} + \frac{۱۱}{۱۲}}} = \frac{۱۲}{\sqrt{\frac{۲۱}{۱۲}}}$$

$$t = \frac{۱۲}{\sqrt{۱,۷۵}} = \frac{۱۲}{۱,۳۲} \Rightarrow t = ۹,۰۹$$

$$t = ۱,۹۶ < t = ۹,۰۹$$

✓ H_1 = بین رضایت شغلی زنان و مردان با یکدیگر تفاوت وجود دارد

✗ H_0 = بین رضایت شغلی زنان و مردان با یکدیگر تفاوت وجود ندارد

با توجه به بزرگتر بودن t محاسبه شده ($t = ۹,۰۹$) از t ملاک یا جدول ($t = ۱,۹۶$) می توان نتیجه گرفت که تفاوت معناداری بین رضایت شغلی زنان و مردان وجود دارد. بنابراین فرض تحقیق (H_1) تایید و فرض خلاف (H_0) رد می شود هم چنین با توجه به بزرگتر بودن میانگین رضایت شغلی زنان از مردان می توان نتیجه گرفت که کارکنان زن بیش از کارکنان مرد از رضایت شغلی برخوردار هستند

P4PCO

Scanned with CamScanner

Subject:
Date

آزمون t گروه های همتا (هم تراز، جفت شده، هم بسته): این آزمون زمانی مورد استفاده قرار می‌گیرد که بخواهیم نمرات یک گروه را در ۲ بازه زمانی مختلف مورد بررسی قرار دهیم به عبارت دیگر نمرات ما در پیش از ارائه متغیر مستقل و پس از آن در دسترس باشد. برای بدست آوردن t همتا از فرمول زیر استفاده می کنیم

$$t = \frac{|\varepsilon D|}{\sqrt{\frac{n\varepsilon D^2 - (\varepsilon D)^2}{n-1}}}$$

پژوهشگری قصد دارد به بررسی تاثیر آموزش مشارکتی بر یادگیری دانش آموزان بپردازد. برای این منظور یک گروه ۱۱ نفره از دانش آموزان را به عنوان نمونه انتخاب نموده و میزان یادگیری آن ها را قبل از استفاده از آموزش مشارکتی اندازه گیری نموده سپس برای مدت ۲ ماه از آموزش مشارکتی در کلاس درس استفاده نمود. پس از پایان دوره آموزش مجدد آموزان یادگیری دانش آموزان را اندازه گیری نمود. نتایج جدول زیر بدست آمده است

قبل از آموزش	بعد آموزش	D	D^2
۱۴	۱۸	-۴	۱۶
۱۲	۱۷	-۵	۲۵
۸	۱۳	-۵	۲۵
۱۵	۱۵	۰	۰
۱۳	۱۱	+۲	۴
۷	۱۴	-۷	۴۹
۱۰	۱۳	-۳	۹
۱۲	۱۵	-۳	۹
۱۸	۱۸	۰	۰
۱۱	۱۴	-۳	۹
۹	۱۵	-۶	۳۶
		-۳۴	۱۸۲

$$t = \frac{|\varepsilon D|}{\sqrt{\frac{n\varepsilon D^2 - (\varepsilon D)^2}{n-1}}}$$

$$t = \frac{|-34|}{\sqrt{\frac{(11\times 182) - (-34)^2}{11-1}}}$$

$$t = \frac{34}{\sqrt{\frac{858}{10}}} = \frac{+34}{\sqrt{85,8}}$$

$$t = \frac{+34}{9,19} = +3,69$$

$t_{\text{جدول}} = 2,58 < t_{\text{محاسبه شده}} = 3,69$

پس یادگیری دانش آموزان قبل از آموزش با یادگیری بعد از آموزش تفاوتی وجود دارد $\Rightarrow H_1$ ✓

پس یادگیری دانش آموزان قبل از آموزش با یادگیری بعد از آموزش تفاوتی وجود ندارد $\Rightarrow H_0$ ✗

فرض تحقیق تایید و فرض خلاف رد می شود.

PqPCO

Scanned with CamScanner

تحلیل واریانس یک طرفه ← ANOVA F

هرگاه بخواهیم به مقایسه میانگین در بین بیش از ۲ گروه بپردازیم می توان از این آزمون استفاده کرد. این آزمون در زمره ی آزمون های پارامتریک قرار می گیرد یعنی متغیر کمی می بایست نرمال باشد. معادل ناپارامتریک آزمون های ارائه شده به شرح زیر می باشد:

t تک نمونه ای - تک گروهی ← خیدو - کای اسکوئر - کایدو

t مستقل ← U ← یو، من ویتنی، ویتنی

t همبسته (همتا) ← ویل کاکسون

F ← کروسکال والیس

before و after داشتن ← ویتنی

پارامتریک	ناپارامتریک
قوی	ضعیف
شربت خارجی	شربت ایرانی

خیدو: این آزمون را می توان معادل ناپارامتریک t تک گروهی دانست زمانی مورد استفاده قرار می گیرد که بخواهیم صرفا به وجود یا عدم وجود تفاوت رابطه در بین دسته ها و فراوانی ها بپردازیم. این آزمون بر اساس فراوانی مشاهده شده و فراوانی مورد انتظار محاسبه می شود و با استفاده از فرمول زیر محاسبه می شود ∴ فراوانی مشاهده شده

$$\chi^2 = \varepsilon \frac{(F_o - F_e)^2}{F_e}$$

χ^2 ← خیدو

F_e ← فراوانی مورد انتظار

مدیر کارخانه ای قصد دارد به ارزیابی نظرات کارکنان خود نسبت به نحوه مدیریت بپردازد برای این منظور ۱۰۰ نفر از کارکنان را جهت شرکت در این ارزیابی انتخاب می نماید پس از بررسی های صورت گرفته مشخص شد ۱۰ کاملا ناراضی و ۱۵ نفر ناراضی و ۳۵ نفر راضی و ۴۰ نفر نیز کاملا از عملکرد مدیر راضی و رضایت کامل را دارند با اطمینان ۹۵ درصد به بررسی وجود تفاوت معنادار در بین نظرات کارکنان بپردازید

χ^2 خیدو محاسبه = ۳,۵۸

$N = 100$

$F_e = \frac{N}{K}$

$F_e = \frac{100}{4} = 25$

Scanned with CamScanner

Subject:
Date:

	کاملا راضی	راضی	ناراضی	کاملا ناراضی
F_o	۴۰	۳۵	۱۵	۱۰
F_c	۲۵	۲۵	۲۵	۲۵
$(F_o - F_e)$	+۱۵	+۱۰	-۱۰	-۱۵
$(F_o - F_e)^2$	۲۲۵	۱۰۰	۱۰۰	۲۲۵
$\frac{(F_o - F_e)^2}{F_e}$	۹	۴	۴	۹

۶

$\chi^2_{\text{محاسبه شده}} = ۲۶ > \chi^2_{\text{ملاک}} = ۳,۵۸$

این مرحله تدوین فرضیه است

✓ H_1: کارکنان از عملکرد مدیر رضایت دارند

✗ H_0: کارکنان از عملکرد مدیر رضایت ندارند

باتوجه به بزرگتر بودن مقدار خیدو محاسبه شده از مقدار خیدو جدول یا ملاک می‌توان نتیجه گرفت که بین فراوانی های مشاهده شده و فراوانی های مورد انتظار تفاوت وجود دارد به عبارت دیگر H_1 یا فرض تحقیق تایید و H_0 رد می‌شود یعنی با اطمینان ۹۵ درصد کارکنان از عملکرد مدیر رضایت دارند

مدیر کارخانه‌ای قصد دارد برای ارزیابی نظرات کارکنان خود نسبت به نحوه مدیریت بپردازد برای این منظور ۲۰۰ نفر از کارکنان جهت شرکت در این ارزیابی انتخاب می‌نماید پس از بررسی های مشخص شده ۴۰ نفر در بخش مالی و ۶۰ نفر در بخش اداری و ۲۰ نفر در فروش و ۳۰ نفر در خدمات و ۵۰ نفر در بخش تولید هستند

$\chi^2 = ۳,۵۸$ خیدو ملاک

	تولید	خدمات	فروش	اداری	مالی
F_o	۵۰	۳۰	۲۰	۶۰	۴۰
F_e	۴۰	۴۰	۴۰	۴۰	۴۰
$(F_o - F_e)$	۱۰	-۱۰	-۲۰	۲۰	۰
$(F_o - F_e)^2$	۱۰۰	۱۰۰	۴۰۰	۴۰۰	۰
$\frac{(F_o - F_e)^2}{F_e}$	۲,۵	۲,۵	۱۰	۱۰	۰

۴ = ۲۵

$F_e = \frac{۸}{۲}$

$F_e = \frac{۲۰۰}{۵} = ۴۰$

تفاوت معناداری وجود دارد زیرا خیدو محاسبه شده از خیدو ملاک بیشتر است.

Scanned with CamScanner

نرم افزار spss

S ↓ علوم — S ↓ اجتماعی — P ↓ پردازنده — S ↓ آمار

Variable view = name ↓ نام — value ↓ ارزش ها — measure ↓ اندازه گیری — lable ↓ برچسب زدن کدها

data view = فقط اعداد را شامل می شود یعنی هرستون یک متغیر داریم

ساختن قالب خام متغیر: ۱- یک نگاه به متغیرهای اندازیم و نوع مقیاس هرمتغیر را سریعا تعیین می کنیم

۲- کدگذاری متغیرهای کیفی

☑ درحین کدگذاری متغیرهای کیفی باید این نکته را مورد توجه قرار داد که اگر مقیاس متغیری رتبه ای باشد کمترین کد به پایین ترین ارزش و بیشترین کد به بالاترین ارزش اختصاص یابد

۳- نام گذاری متغیرها در ستون name صفحه Variable view

☑ سعی کنیم در نام گذاری متغیر از کلمات اختصار استفاده کنیم و به جای نوشتن زبان فارسی از حروف انگلیسی استفاده کنیم

☑ کوچک ترین واحد کامپیوتر وِیت است

۴- درصورت نیاز به ارائه توضیحات اضافی برای یک متغیرمی توان در lable آن متغیر آن توضیحات را ارائه نمود

۵- تعریف کد های درنظر گرفته شده برای متغیرهای کیفی درقسمت Value هرمتغیر

۶- تعیین مقیاس هرمتغیردرستون Measure

ordinal = رتبه ای - یکان

Normnal = اسمی

scale = درجه بندی

ما ۷ متغیر و ۷ ستون داریم

Var = متغیر data = داده

منظور از رتبه ای یعنی بهترین و بدترین مقیاس است.

PqPCO

Scanned with CamScanner

① جنسیت: زن ☐ مرد ☐ کیفی ← اسمی ✓ (۲) / رتبه‌ای

② شغل: کارمند ☐ خانه‌دار ☐ محصل ☐ (۱ ۲ ۳)

③ سن: ــــ کمی ← فاصله‌ای / نسبی

④ تاهل: مجرد ☐ متاهل ☐ کیفی ← اسمی ✓ (۲) / رتبه‌ای

⑤ رشته کارشناسی: علوم انسانی ☐ تجربی ☐ ریاضی ☐ هنر ☐ (۱ ۲ ۳ ۴) کیفی ← اسمی ✓ / رتبه‌ای

⑥ نوع دانشگاه: دولتی ☐ غیردولتی ☐ کیفی ← اسمی ✓ (۲) / رتبه‌ای

⑦ معدل کارشناسی: ــــ کمی

Analyze = بخش آماری

ابتدا از بخش Analyze به قسمت Descriptive statis رفته و گزینه Frequancy و قسمت charts را انتخاب می‌کنیم. یک پنجره باز می‌شود که با توجه به نکات زیر باید انتخاب کنیم (charts = نمودار)

Histogram ← نمودارهای به هم چسبیده ـ برای متغیرهای کمی و عددی کاربرد دارد

par chart ← نمودار دایره‌ای

bar chart ← نمودار میله‌ای

☒ اگر نمودار مورد بررسی ما کیفی باشد می‌توان نمودار میله‌ای و نمودار دایره‌ای استفاده کنیم که معمولا اگر مقیاس ما اسمی بود نمودار دایره‌ای و اگر مقیاس ما رتبه‌ای بود از نمودار میله‌ای استفاده می‌کنیم

statistics ← برای متغیر کمی کاربرد دارد (میانه ـ مد ـ واریانس)

mean = معنی داشتن median = میانه mode = نما

sum = Σ ـ مجموع S.E mean = خطای انحراف از میانگین (آزمون t تک گروهی)

st deviation = انحراف Quartiles = چارک

تحلیل یا پایان نامه ← تحلیل توصیفی / تحلیل استنباطی

ضرایب هم بستگی ← پیرسون / اسپیرمن

آزمون های مقایسه‌ای ← t تک گروهی / t مستقل / t همتا

F (SPSS) ← محاسبه‌ی زیاد

Scanned with CamScanner

ابتدا به قسمت Analays رفته و قسمت Correlate را انتخاب و بخش bivariate را انتخاب می کنیم (ضرایب پیرسون و ضرایب اسپیرمن در این بخش هستند)

مثال امتحانی) پژوهشگری قصد دارد ارتباط بین سن و معدل دانش آموزان را بررسی بپردازد پس از بررسی صورت گرفته و جمع آوری اطلاعات نتایج جدول زیر به دست آمد مطلوب است به تدوین فرضیه و تفسیر آن متناسب با اعداد جدول بپردازد

Sig = خطا

آیا بین سن و معدل دانش آموزان رابطه وجود دارد؟

H_1 بین سن و معدل دانش آموزان رابطه وجود دارد.

H_0 بین سن و معدل دانش آموزان رابطه وجود ندارد.

Sig < ۰/۰۵ ← رابطه وجود ندارد

Sig > ۰/۰۵ رابطه وجود دارد

باتوجه به اعداد جدول فوق (به Sig نگاه می کنیم) و مقدار سطح معناداری آزمون Sig = ۰/۱۹ است و بزرگتر بودن آن از ۰/۰۵ می توان نتیجه گرفت به آزمون معنادار نمی باشد بنابراین فرض تحقیق رد و فرض خلاف تایید می شود

☒ اگر یک متغیر دارای مقیاس رتبه ای و متغیر بعدی اسمی نباشد شرایط استفاده از اسپیرمن را دارد. چون Sig = ۰/۱۹ شده پس فرض خلاف تایید و فرض تحقیق رد می شود و آزمون معنادار نیست

پژوهشگری قصد دارد ادعای مدیر آموزش دانشکده ای را مبنی بر آنکه معدل دانشجویان آن دانشکده ۱۳۰ خواهد بود بپردازد پس از جمع آوری اطلاعات مورد نیاز اعداد جدول زیر به دست آمد مطلوب است به مقایسه میانگین افراد با میانگین جامعه بپردازید؟

☒ اگر در مسئله کلمه مقایسه آمد یعنی مسئله مربوط به آزمون های t است و در تفسیر باید کلمه تفاوت آورده شود در غیر این صورت در تفسیر مسئله از کلمه رابطه وجود دارد استفاده می کنیم

Scanned with CamScanner

آیا بین میانگین معدل دانشجویان با میانگین جامعه تفاوت وجود دارد؟
H_1 بین میانگین معدل دانشجویان با میانگین جامعه تفاوت وجود دارد.
H_0 بین میانگین معدل دانشجویان با میانگین جامعه تفاوت وجود ندارد.

پس چون sig است و باتوجه به مقدار سطح معناداری آزمون و نیز جامعه بودن آن از نظر می توان نتیجه گرفت که آزمون معنادار می باشد پس این فرض تایید و فرض خلاف رد می شود.

آزمون ANOVA (تحلیل واریانس): ابتدا به قسمت ANAlays رفته و جنس Independent رفته. حال متغیر هایی که داریم به قسمت سمت راست انتقال داده و continue را زده سپس محاسبه جدول طاهر می شود. اگر به صورت دستی بخواهیم محاسبه کنیم خیلی طولانی هست و نرم افزار spss این کار را انجام داده است

در این محاسبه ANOVA ← مقیاس های کمی (عددی) در بالا و مقیاس های کیفی در پایین قرار دارند.

باتوجه به نمودار sig آزمون leven ۰/۵۴۷ است و بزرگتر بودن آن از ۰/۰۵ می توان نتیجه گرفت که اعداد آزمون t در سطح برابری واریانس ها (ردیف بالا) مورد تایید قرار می گیرد بنابراین سطح معناداری آزمون t برابر با ۰/۶۰ خواهد بود. باتوجه به بزرگتر بودن معناداری از ۰/۰۵ می توان نتیجه گرفت که آزمون t ارائه شده معنادار نخواهد بود یعنی بین میانگین معدل دانشجویان دختر با پسر تفاوتی مشاهده نمی شود بنابراین فرض تحقیق رد و فرض خلاف تایید می شود.

(چون sig بیشتر از ۰/۰۵ است پس ردیف بالا را تفسیر می کنیم اگر sig کمتر از ۰/۰۵ باشد ردیف پایین تفسیر می شود)

گاه اگر بخواهیم به متغیر کمی نرمال را در بیش از ۲ گروه در مورد مقیاس قرار دهیم می توان از آزمون ANOVA (تحلیل واریانس) استفاده کرد.

این جدول ها و فرمول ها در امتحان داده می شود.

Scanned with CamScanner