

نمونه سوالات امتحان آمار توصیفی و زیستی در ترمهای قبل - (بعضاً نمونه تکراری هستند)

1) جمعیت چیست نمونه چیست (10نمره)

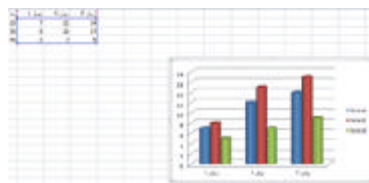
جمعیت: مجموعه تمام عناصری که دارای یک یا چند ویژگی مشترک باشند (مثلا در یک دانشگاه با 400 دانشجو اعداد وزن تمام 400 نفر دانشجو را داشته باشیم)

نمونه: بخشی از جمعیت میباشد (مثلا در یک دانشگاه با 400 دانشجو اعداد وزن 80 نفر دانشجو را داشته باشیم)

2) انواع مقیاسها را نام برده و توضیح ببرید (10نمره)

مقیاس اسمی (NOMINAL) - کد گذاری عددی - اعداد نسبت داده شده ارجحیت و برتری ندارند (مثل جنسیت - زن 2 مرد 1)
 مقیاس رتبه‌ای (ORDINAL) (ترتیبی) - با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند ولی تناسب ندارند (مثل پرسشنامه لینکرت شامل موافق 3 - بی نظر 2 - مخالف 1)
 مقیاس نسبتی (SCALE) (وزنی) با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند و تناسب هم دارند (مثل سن و مثل وزن - مثلا وزن احمد 35 و وزن عباس 70 کیلوگرم یعنی عباس بیشتر از احمد است و به نسبت دو برابر هم میباشد)
 مقیاس فاصله‌ای - مثلا به یک بیمار بگوییم اگر حداکثر درد مثل یک خط کش عدد 100 باشد درد شما چقدر است و بیمار بگوید 65

3) دو مورد از انواع نمودارها نام ببرید و هر یک نیم خط توضیح دهید و شکل نمودار را ترسیم کنید (10نمره)



نمودار میله ای Bar Chart: برای مشاهده فراوانی داده ها و مقایسه داده ها نسبت به هم

نمودار دایره ای Pie Chart: برای مشاهده سهم هر مورد از داده ها

نمودار هیستوگرام Histogram: برای مشاهده نحوه توزیع داده ها بصورت پیوسته

نمودار خطی Line Chart: برای نشان دادن رابطه بین دوسری اعداد مثلا خط رگرسیون



نمودار پراکنندگی Scatter Plot: برای نشان دادن پراکنندگی بین دوسری اعداد مثلا در رگرسیون



4) نمرات دانشجویان بشرح زیر است داده‌ها را در جدول زیر مرتب کنید. الف) میانه داده ها چه مقدار است؟
 (چرا) ب) مد(نما) داده ها چه مقدار است. - ج) دهک چهارم داده‌های را معلوم کنید - د) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید. ه) ضریب تغییرات را بدست آورید و با تکمیل جدول، فراوانی نسبی و فراوانی نسبی تجمعی را محاسبه کنید و با یک مثال نسبی ها را توضیح دهید .. نمرات: 12 و 11 و 18 و 13 و 11 و 13 و 12 و 10 و 13 (40نمره)

| | | | | | |
|-----------|--|--|--|--|--|
| نمره = x | | | | | |
| تعداد = f | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

| | | | | | |
|--------|----|----|----|----|----|
| x داده | 10 | 11 | 12 | 13 | 18 |
|--------|----|----|----|----|----|

| | | | | | |
|--|----------|----------|----------|----------|------------|
| فراوانی f | 1 | 2 | 2 | 4 | 1 |
| فراوانی تجمعی F | 1 | 3 | 5 | 9 | 10 |
| فراوانی نسبی $r = \frac{f_i}{\sum f_i}$ | 1/10=10% | 2/10=20% | 2/10=40% | 4/10=20% | 1/10=10% |
| فراوانی نسبی تجمعی $R = \frac{\sum f_{i_i}}{\sum f_i}$ | 1/0=10% | 3/10=30% | 5/10=70% | 9/10=90% | 10/10=100% |

مد داده ای بمقدار 13 میباشد زیرا بیشترین تعداد را دارد

M=13

میانۀ داده ها

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 2 + 4 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=9 \rightarrow m=13$$

یعنی نصف داده ها 13 یا کمتر از 13 هستند
یا

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 2 + 4 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=5 \rightarrow m=12$$

یعنی نصف داده ها 12 یا کمتر از 12 هستند

دهک 4

$$Q \sum f_i = \frac{4}{10}(1 + 2 + 4 + 2 + 1) = \frac{4}{10} = 4 \rightarrow \text{در سطر } F_i \rightarrow 4+ \rightarrow F_i=5 \rightarrow 12$$

0.4 داده ها 12 یا کمتر از 12 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(10 * 1) + (11 * 2) + (12 * 2) + (13 * 4) + (18 * 1)}{1 + 2 + 2 + 4 + 1} = \frac{126}{10} = 12.6$$

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f_i}{\sum f_i}$$

$$\sigma^2 = \frac{(10 - 12.6)^2 * 1 + (11 - 12.6)^2 * 2 + (12 - 12.6)^2 * 2 + (13 - 12.6)^2 * 4 + (18 - 12.6)^2 * 1}{1 + 2 + 4 + 2 + 1} = 4.24 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.24} = 2.05 \text{ انحراف معیار}$$

ضریب تغییرات

$$\rho = \frac{\sigma}{\mu} = \frac{2.05}{12.6} = 16\%$$

یعنی پراکندگی داده‌ها معمولی است

همیشه $0 \leq \rho \leq 1$ است

اگر ρ صفر شود یعنی داده‌ها اصلاً پراکنده نیستند داده‌ها بسیار بهم نزدیک هستند

اگر ρ یک شود یعنی داده‌ها خیلی پراکنده هستند

جدول در بالا تکمیل شد

نسبی r یعنی اینکه مثلاً از داده به مقدار 13 بتعداد 4 تا داریم یعنی 40٪ داده ها 13 هستند
 نسبی تجمعی R یعنی اینکه مثلاً از داده به مقدار 13 و کمتر از 13 بتعداد 9 تا داریم یعنی 09٪ داده ها 13 یا کمتر از 13 هستند

5) در یک منطقه میزان مصرف برق در سالهای گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سالهای بعد چقدر میباشد (20نمره)

| | | | | | | | |
|---|------|------|------|------|------|------|------|
| X | 1394 | 1395 | 1396 | 1397 | 1398 | 1399 | 1400 |
| Y | 2 | 2 | 2 | 3 | 4 | ؟ | ؟ |

سال مبنا را سال 1396 در نظر میگیریم و همه X ها را از 1396 کم میکنیم

| | | | | | | | |
|-------|----|----|---|----|----|----|----|
| X=سال | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| مصرف | 2 | 2 | 2 | 3 | 4 | ?? | ?? |

$$y = a + bx, \quad b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}, \quad \bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$= \frac{(-2 * 2) + (-1 * 2) + (0 * 2) + (1 * 3) + (2 * 4) - \frac{(-2 - 1 + 0 + 1 + 2) * (2 + 2 + 2 + 3 + 4)}{5}}{(-2)^2 + (-1)^2 + 0^2 + (1)^2 + (2)^2 - \frac{(-2 - 1 + 0 + 1 + 2)^2}{5}}$$

$$b = \frac{-4 - 2 + 0 + 3 + 8 - \frac{0 * 16}{5}}{4 + 1 + 0 + 1 + 4 - \frac{0}{3}} \quad b = \frac{5}{10} = 0.5$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-2 - 1 + 0 + 1 + 2}{5} \quad \bar{x} = 0$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{2 + 2 + 2 + 3 + 4}{5} \quad \bar{y} = \frac{13}{5} = 2.6$$

$$\bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x} \rightarrow a = 2.6 - 0.8 * 0 \rightarrow a = 2.6 \rightarrow y = a + bx \rightarrow$$

$$y = 2.6 + 0.5x$$

$$1399 - 1396 = 3 \rightarrow y = 2.6 + 0.5x \rightarrow y = 2.6 + (0.5 * 3) = 2.6 + 1.5 = 4.1$$

پس برای سال 1399 پیش بینی به مقدار 4.1 میشود

$$1400 - 1396 = 4 \rightarrow y = 2.6 + 0.5x \rightarrow y = 2.6 + (0.5 * 4) = 2.6 + 2 = 4.6$$

پس برای سال 1400 پیش بینی به مقدار 4.6 میشود

6) جدول سوال زیر را قبلاً حل شده است و میانگین 14.3 و واریانس 14.8 میباشد. مجدد حل نکنید

| | | | | | |
|-------|----|----|----|----|----|
| X_i | 10 | 12 | 14 | 15 | 17 |
| f_i | 1 | 3 | 2 | 2 | 2 |

الف) انحراف معیار را حساب کنید

ج) میانگین با انحراف معیار جمع و تفریق کنید و یکبار دیگر میانگین با دو برابر انحراف معیار جمع و تفریق کنید

و برای داده فوق که نرمال است هر دو جواب را کامل توضیح دهید . د) جدول زیر شباهت زیادی به جدول بالا در همین سوال دارد با استفاده از جواب قسمت بالا میانگین و انحراف معیار جدول زیر را بدون حل و بشکل سریع بدست آورید (10 نمره)

| | | | | | |
|-------|----|----|----|----|----|
| x_i | 20 | 22 | 24 | 25 | 27 |
| f_i | 1 | 3 | 2 | 2 | 2 |

در داده های فوق هر یک از داده جدید باندازه 60 واحد بیش از داده قدیم است در خصوص x ها خود مسئله عنوان کرده که

$$\bar{x} = 14.3 \quad \sigma_x = \sqrt{14.8} = 3.84$$

طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 10$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 10 = 14.3 + 10 \rightarrow \bar{y} = 24.3$$

$$\sigma_y^2 = \sigma_x^2 \rightarrow$$

$$\sigma_y = \sigma_x = \sqrt{14.8} = 3.84 \rightarrow \sigma_y = 3.84$$

چون داده های جدید همگی 10 واحد اضافه شده اند پس میانگین داده های جدید 10 واحد بیشتر از میانگین داده های قبلی میشود و انحراف معیار داده های جدید با انحراف معیار داده های قدیم هیچ فرقی نمیکند

=====

7) حقوق ده نفر بشرح زیر است دادهها را در جدول زیر مرتب کنید. a) مد(نما) داده ها چه مقدار است (چرا) b). میانه داده ها چه مقدار است؟ c) صدک 72 دادههای را معلوم کنید - d) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید. e) با تکمیل جدول، فراوانی نسبی و فراوانی نسبی تجمعی را محاسبه کنید و با یک مثال نسبی ها را توضیح دهید. f) یکبار میانگین با انحراف معیار جمع و تفریق کنید و یکبار میانگین با دو برابر انحراف معیار جمع و تفریق کنید و هر دو را توضیح دهید. g) ضریب تغییرات را بدست آورید h) فرمول واریانس بدون توان 2 آن بنویسید و حاصل را بدست آورید اعداد حقوق: 5 و 3 و 2 و 9 و 5 و 2 و 3 و 3 و 1 و 3

حل:

| | | | | | |
|--|----------|----------|----------|----------|------------|
| داده x | 1 | 2 | 3 | 5 | 9 |
| فراوانی f | 1 | 2 | 4 | 2 | 1 |
| فراوانی تجمعی F | 1 | 3 | 7 | 9 | 10 |
| فراوانی نسبی $r = \frac{f_i}{\sum f_i}$ | 1/10=10% | 2/10=20% | 4/10=40% | 2/10=20% | 1/10=10% |
| فراوانی نسبی تجمعی $R = \frac{\sum f_{i_i}}{\sum f_i}$ | 1/0=10% | 3/10=30% | 7/10=70% | 9/10=90% | 10/10=100% |

=====

مد داده ای بمقدار 3 میباشد زیرا بیشترین تعداد را دارد

$$M=3$$

=====

میانه داده ها

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 4 + 2 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } Fi \rightarrow 5+ \rightarrow Fi=7 \rightarrow m=3$$

یعنی نصف داده ها 3 یا کمتر از 3 هستند

صدک 72

$$Q \sum f_i = \frac{25}{100}(1 + 2 + 4 + 2 + 1) = \frac{72}{100} = 7.2 \rightarrow \text{در سطر } Fi \rightarrow 7.2+ \rightarrow Fi=9 \rightarrow 5$$

72٪ داده ها 5 یا کمتر از 5 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(1 * 1) + (2 * 2) + (3 * 4) + (5 * 2) + (9 * 1)}{1 + 2 + 4 + 2 + 1} = \frac{36}{10} = 3.6$$

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f_i}{\sum f_i}$$

$$\sigma^2 = \frac{(1 - 3.6)^2 * 1 + (2 - 3.6)^2 * 2 + (3 - 3.6)^2 * 4 + (5 - 3.6)^2 * 2 + (9 - 3.6)^2 * 1}{1 + 2 + 4 + 2 + 1} = \frac{6.76 + 5.12 + 1.44 + 3.92 + 29.16}{1 + 2 + 4 + 2 + 1} = 4.6 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.6} = 2.14 \text{ انحراف معیار}$$

$$\mu \pm \sigma \quad \mu \pm 2\sigma$$

$$\mu \pm \sigma = 3.6 \pm 2.14 = 1.46 \text{ تا } 5.74$$

یعنی بیش از 68٪ داده ها بین 1.46 تا 5.74 میباشند

$$\mu \pm 2\sigma = 3.6 \pm (2 * 2.14) = 3.6 \pm 4.28 = -0.69 \text{ تا } 7.89$$

یعنی بیش از 96٪ داده ها بین -0.69 تا 7.89 میباشند

ضریب تغییرات

$$\rho = \frac{\sigma}{\mu} = \frac{2.14}{3.6} = 0.46 = 46\%$$

یعنی پراکندگی داده ها معمولی است

همیشه $0 \leq \rho \leq 1$ است

اگر ρ صفر شود یعنی داده ها اصلا پراکنده نیستند داده ها بسیار بهم نزدیک هستند

اگر ρ یک شود یعنی داده ها خیلی پراکنده هستند

$$\frac{\sum (x_i - \bar{x}) * f_i}{\sum f_i} = \frac{(1 - 3.6) * 1 + (2 - 3.6) * 2 + (3 - 3.6) * 4 + (5 - 3.6) * 2 + (9 - 3.6) * 1}{1 + 2 + 4 + 2 + 1}$$

$$\frac{-2.6 - 3.2 - 2.4 + 2.8 + 5.4}{1 + 2 + 4 + 2 + 1} = \frac{0}{10} = 0$$

جدول در بالا تکمیل شد

نسبی r یعنی اینکه مثلا از داده به مقدار 3 بتعداد 4 تا داریم یعنی 40٪ داده ها 3 هستند

نسبی تجمعی R یعنی اینکه مثلاً از داده به مقدار 3 و کمتر از 3 بتعداد 7 تا داریم یعنی 70٪ داده ها 3 یا کمتر از 3 هستند

(8) در یک منطقه میزان مبتلایان به ازدیاد قندخون در چند سال گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال 1399 چقدر میباشد

| | | | | | | |
|-------|------|------|------|------|------|------|
| سال=x | 1394 | 1395 | 1396 | 1397 | 1398 | 1399 |
| مصرف | 2 | 2 | 3 | 4 | 5 | ?? |

سال مبنا را سال 1396 در نظر میگیریم و همه X ها را از 1396 کم میکنیم

| | | | | | | |
|-------|----|----|---|----|----|----|
| سال=x | -2 | -1 | 0 | +1 | +2 | +3 |
| مصرف | 2 | 2 | 3 | 4 | 5 | ?? |

$$y = a + bx, \quad b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}, \quad \bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$= \frac{(-2 * 2) + (-1 * 2) + (0 * 3) + (1 * 4) + (2 * 5) - \frac{(-2 - 1 + 0 + 1 + 2) * (2 + 2 + 3 + 4 + 5)}{5}}{(-2)^2 + (-1)^2 + 0^2 + (1)^2 + (2)^2 - \frac{(-2 - 1 + 0 + 1 + 2)^2}{5}}$$

$$b = \frac{-4 - 2 + 0 + 4 + 10 - \frac{0 * 16}{5}}{4 + 1 + 0 + 1 + 4 - \frac{0}{5}} \quad b = \frac{8}{10} = 0.8$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-2 - 1 + 0 + 1 + 2}{5} \quad \bar{x} = 0$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{2 + 2 + 3 + 4 + 5}{5} \quad \bar{y} = \frac{16}{5} = 3.2$$

$$\bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x} \rightarrow a = 3.2 - 0.8 * 0 \rightarrow a = 3.2 \rightarrow y = a + bx \rightarrow$$

$$y = 3.2 + 0.8x$$

$$1399 - 1396 = 3 \rightarrow y = 3.2 + 0.8x \rightarrow y = 3.2 + (0.8 * 3) = 3.2 + 2.4 = 5.6$$

پس برای سال 98 پیش بینی به مقدار 5.6 میشود

(9) در جدول زیر در خصوص X میانگین 4.7 واریانس 5.7 انحراف معیار 2.38 قبلاً خودمان محاسبه کرده ایم اصلاً مجدداً محاسبه نکنید. در خصوص داده های Y که با یک نسبت به X تغییر یافته اند جوابهای میانگین واریانس و انحراف معیار Y بدون استفاده از فرمولهای اصلی و سریع از طریق جوابهای X بنویسید

| | | | | | |
|-----------|----|----|----|----|----|
| x = شماره | 2 | 4 | 5 | 6 | 7 |
| f = تعداد | 20 | 10 | 40 | 20 | 10 |

| | | | | | |
|-----------|----|----|----|----|----|
| y = وزن | 62 | 64 | 65 | 66 | 67 |
| f = تعداد | 20 | 10 | 40 | 20 | 10 |

قانونی داریم که اگر داده ها به یک مقدار یکسان اضافه شوند میانگین هم به همان مقدار اضافه میشود ولی واریانس و انحراف معیار ثابت میماند

در داده های فوق هر یک از داده جدید باندازه 60 واحد بیش از داده قدیم است
در خصوص x ها خود مسئله عنوان کرده که

$$\bar{x} = 4.7 \quad \sigma_x = 2.38$$

طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 60$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 60 = 4.7 + 60 \rightarrow \bar{y} = 64.7$$

$$\sigma_y^2 = \sigma_x^2 \rightarrow \sigma_y = 5.7$$

$$\sigma_y = \sigma_x = \sqrt{5.7} = 2.38 \rightarrow \sigma_y = 2.38$$

چون داده های جدید همگی 60 واحد اضافه شده اند پس میانگین داده های جدید 60 واحد بیشتر از میانگین داده های قبلی میشود و انحراف معیار داده های جدید با انحراف معیار داده های قدیم هیچ فرقی نمیکند

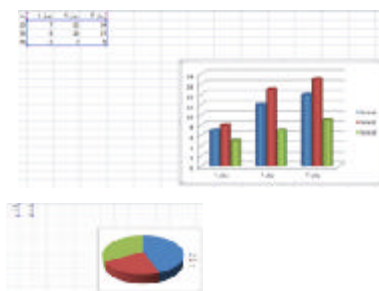
10) آمار توصیفی و آمار استنباطی را توضیح دهید

آمار توصیفی، مجموعه ای از روش هایی است که برای سازمان دهی، خلاصه کردن، تهیه جدول، رسم نمودار، توصیف و تفسیر داده های جمع آوری شده از نمونه آماری به کار گرفته می شود - مراحل آن : خلاصه کردن داده ها و توصیف الگوی کلی - فشرده کردن داده ها در قالب جدول های آماری - نمایش آن ها به وسیله نمودار - محاسبه شاخص های آماری
آمار استنباطی مشخص می کند که آیا الگوها و فرآیندهای کشف شده در نمونه، در جامعه آماری هم کاربرد دارد یا خیر. بنابراین، آمار استنباطی راجع به ویژگی ها و پارامترهای مربوط به جامعه آماری تحقیق و کیفیت ارتباط بین مفاهیم و متغیرها می باشد. بدین ترتیب، می توان گفت که از آمار استنباطی در تجزیه و تحلیل مقایسه ای و رابطه ای (علی - همبستگی) استفاده می شود.

11) انواع مقیاسها را نام برده و توضیح ببرید

مقیاس اسمی (NOMINAL) - کد گذاری عددی - اعداد نسبت داده شده ارجحیت و برتری ندارند (مثل جنسیت - زن 2 مرد 1)
مقیاس رتبه ای (ORDINAL) (ترتیبی) - با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند ولی تناسب ندارند (مثل پرسشنامه لیکرت شامل موافق 3 - بی نظر 2 - مخالف 1)
مقیاس نسبتی (SCALE) (وزنی) با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند و تناسب هم دارند (مثل سن و مثل وزن - مثلا وزن احمد 35 و وزن عباس 70 کیلوگرم یعنی عباس بیشتر از احمد است و به نسبت دو برابر هم میباشد)
مقیاس فاصله ای - مثلا به یک بیمار بگوییم اگر حداکثر درد مثل یک خط کش عدد 100 باشد درد شما چقدر است و بیمار بگوید 65

12) دو مورد از انواع نمودارها نام ببرید و هر یک نیم خط توضیح دهید و شکل نمودار را ترسیم کنید



نمودار میله ای Bar Chart: برای مشاهده فراوانی داده ها و مقایسه داده ها نسبت به هم

نمودار دایره ای Pie Chart: برای مشاهده سهم هر مورد از داده ها

نمودار هیستوگرام Histogram: برای مشاهده نحوه توزیع داده ها بصورت پیوسته

نمودار خطی Line Chart: برای نشان دادن رابطه بین دوسری اعداد مثلا خط رگرسیون

نمودار پراکندگی Scatter Plot: برای نشان دادن پراکندگی بین دوسری اعداد مثلا در رگرسیون



=====

13) در جدول زیر نمرات 19 نفر از دانشجویان و فراوانی آنها داده شده.

| | | | | | | | | |
|---------|-------|----|----|----|----|----|----|---|
| داده | X_i | 18 | 16 | 14 | 13 | 12 | 11 | 9 |
| فراوانی | f_i | 2 | 4 | 5 | 4 | 2 | 1 | 1 |

الف - جدول را یکبار دیگر با نوشتن جدول مقادیر فراوانی تجمعی و فراوانی نسبی و تجمعی بنویسید

ب - میانه چه عددی است - نما چه عددی است ج - صدک 72 چقدر است

د - میانگین چه مقدار است (دورقم اعشار) ه - واریانس و انحراف معیار چقدر است (دورقم اعشار)

و - این مقادیر را محاسبه و تحلیل کنید $\mu \pm \sigma$ $\mu \pm 2\sigma$

ز - با استفاده از حل موارد فوق در جدول زیر میانگین و انحراف معیار بدون حل کامل محاسبه کنید

| | | | | | | | | |
|---------|-------|----|----|----|----|----|----|----|
| داده | X_i | 20 | 18 | 16 | 15 | 14 | 13 | 11 |
| فراوانی | f_i | 2 | 4 | 5 | 4 | 2 | 1 | 1 |

حل :

الف) داده ها از چپ به راست منظم و مرتب میکنیم

| | | | | | | | | |
|---------------------------------------|-------|----------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|
| داده | X_i | 9 | 11 | 12 | 13 | 14 | 16 | 18 |
| فراوانی | f_i | 1 | 1 | 2 | 4 | 5 | 4 | 2 |
| فراوانی تجمعی | Fi | 1 | 2 | 4 | 8 | 13 | 17 | 19 |
| فراوانی نسبی = $\frac{f_i}{\sum f_i}$ | ri | $\frac{1}{19}$ | $\frac{1}{19}$ | $\frac{2}{19}$ | $\frac{4}{19}$ | $\frac{5}{19}$ | $\frac{4}{19}$ | $\frac{2}{19}$ |
| فراوانی نسبی تجمعی | Ri | $\frac{1}{19}$ | $\frac{2}{19}$ | $\frac{4}{19}$ | $\frac{8}{19}$ | $\frac{13}{19}$ | $\frac{17}{19}$ | $\frac{19}{19}$ |

ب) نما - نما (مود) داده ای که f_i آن از دیگر داده ها دارای فراوانی بیشتر است در اینجا داده ه $M=14$ میباشد
 میانه - باید وسط صف منظم داده ها مشخص کنیم چون داده ها دارای فراوانی است نوشتن تمام آنها (نوشتن تمام 19 داده)
 جدول بزرگتری میخواهد که بهتر است از فرمول استفاده کنیم

$$Q \sum f_i = \frac{1}{2}(1 + 1 + 2 + 4 + 5 + 4 + 2) = \frac{19}{2} = 9.5 \rightarrow \text{در سطر } Fi \rightarrow 9.5+ \rightarrow Fi=13 \rightarrow m=14$$

میانه 14 است

صدک 72

$$Q \sum f_i = \frac{72}{100}(1 + 1 + 2 + 4 + 5 + 4 + 2) = \frac{72 \cdot 19}{100} = 13.68 \rightarrow \text{در سطر } Fi \rightarrow 13.68+ \rightarrow Fi=17$$

$\rightarrow 16$

یعنی 72٪ داده ها 16 یا کمتر از 16 میباشدند

میانگین

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(18 * 2) + (16 * 4) + (14 * 5) + (13 * 4) + (12 * 2) + (11 * 1) + (9 * 1)}{1 + 1 + 2 + 4 + 5 + 4 + 2}$$

$$= \frac{266}{19} = 14$$

واریانس

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f_i}{\sum f_i}$$

$$\sigma^2 = \frac{(9-14)^2 * 1 + (11-14)^2 * 1 + (12-14)^2 * 2 + (13-14)^2 * 4 + (14-14)^2 * 5 + (16-14)^2 * 4 + (18-14)^2 * 2}{1 + 1 + 2 + 4 + 5 + 4 + 2}$$

$$= 4.95$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.95} = 2.22$$

بنابراین میانگین 14 واریانس 4.95 و انحراف معیار 2.22 میباشد
بقیه در صفحه بعد

تحلیل $\mu \pm \sigma$ $\mu \pm 2\sigma$

یعنی بیش از 68٪ داده ها بین 11.78 تا 16.22 میباشند 16.22 تا $\mu \pm \sigma = 14 \pm 2.22 = 11.78$
یعنی بیش از 96٪ داده ها بین 9.56 تا 18.44 میباشند $\mu \pm 2\sigma = 14 \pm (2 * 2.22) = 14 \pm 4.44 = 9.56$ تا 18.44

داده جدید و مقایسه با داده قدیم

| | | | | | | | | |
|---------------|----------------|----|----|----|----|----|----|----|
| x= داده قدیم | X _i | 18 | 16 | 14 | 13 | 12 | 11 | 9 |
| فراوانی | f _i | 2 | 4 | 5 | 4 | 2 | 1 | 1 |
| y = داده جدید | X _i | 20 | 18 | 16 | 15 | 14 | 13 | 11 |

قانونی داریم که اگر داده‌ها به یک مقدار یکسان اضافه شوند میانگین هم بهمان مقدار اضافه میشود ولی واریانس ثابت میماند
در داده‌های فوق هر یک از داده جدید باندازه 2 واحد بیش از داده قدیم است
در خصوص x ها در صفحه قبل بدست آوردیم که

$$\bar{x} = 14 \quad \sigma_x = 2.22$$

طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 2$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 2 = 14 + 2 \rightarrow \bar{y} = 16 \quad \sigma_y = \sigma_x \rightarrow \sigma_y = 2.22$$

میانگین داده‌های جدید دو واحد بیشتر از میانگین داده قدیم است و
واریانس داده جدید مساوی واریانس داده قدیم است

=====

14) حقوق ده نفر بشرح زیر است داده‌ها را در جدول زیر مرتب کنید. (a) مد(نما) داده‌ها چه مقدار است (چرا) (b). میانه داده‌ها چه مقدار است؟ (c) صدک 75 داده‌های را معلوم کنید (d) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید. (e) با تکمیل جدول، فراوانی نسبی و فراوانی نسبی تجمعی را محاسبه کنید و با یک مثال نسبی‌ها را توضیح دهید. (f) یکبار میانگین با انحراف معیار جمع و تفریق کنید و یکبار میانگین با دو برابر انحراف معیار جمع و تفریق کنید و هر دو را توضیح دهید. (g) ضریب تغییرات را بدست آورید (h) فرمول واریانس بدون توان 2 آن بنویسید و حاصل را بدست آورید
اعداد حقوق: 5 و 3 و 2 و 9 و 5 و 2 و 3 و 3 و 1 و 3
حل:

| | | | | | |
|--|----------|----------|----------|----------|------------|
| داده x | 1 | 2 | 3 | 5 | 9 |
| فراوانی f | 1 | 2 | 4 | 2 | 1 |
| فراوانی تجمعی F | 1 | 3 | 7 | 9 | 10 |
| فراوانی نسبی $r = \frac{f_i}{\sum f_i}$ | 1/10=10% | 2/10=20% | 4/10=40% | 2/10=20% | 1/10=10% |
| فراوانی نسبی تجمعی $R = \frac{\sum f_{i_i}}{\sum f_i}$ | 1/0=10% | 3/10=30% | 7/10=70% | 9/10=90% | 10/10=100% |

مد داده ای بمقدار 3 میباشد زیرا بیشترین تعداد را دارد

$$M=3$$

میانۀ داده ها

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 4 + 2 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=7 \rightarrow m=3$$

یعنی نصف داده ها 3 یا کمتر از 3 هستند

صدک 75

$$Q \sum f_i = \frac{75}{100}(1 + 2 + 4 + 2 + 1) = \frac{75}{100} = 7.5 \rightarrow \text{در سطر } F_i \rightarrow 7.5+ \rightarrow F_i= 9 \rightarrow 5$$

75٪ داده ها 5 یا کمتر از 5 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(1 * 1) + (2 * 2) + (3 * 4) + (5 * 2) + (9 * 1)}{1 + 2 + 4 + 2 + 1} = \frac{36}{10} = 3.6$$

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f_i}{\sum f_i}$$

$$\sigma^2 = \frac{(1 - 3.6)^2 * 1 + (2 - 3.6)^2 * 2 + (3 - 3.6)^2 * 4 + (5 - 3.6)^2 * 2 + (9 - 3.6)^2 * 1}{1 + 2 + 4 + 2 + 1} = \frac{6.76 + 5.12 + 1.44 + 3.92 + 29.16}{1 + 2 + 4 + 2 + 1} = 4.6 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.6} = 2.14 \text{ انحراف معیار}$$

$$\mu \pm 2\sigma \quad \mu \pm \sigma \text{ تحلیل}$$

$$\mu \pm \sigma = 3.6 \pm 2.14 = 1.46 \text{ تا } 5.74$$

یعنی بیش از 68٪ داده ها بین 1.46 تا 5.74 میباشدند

$$\mu \pm 2\sigma = 3.6 \pm (2 * 2.14) = 3.6 \pm 4.28 = -0.69 \text{ تا } 7.89$$

یعنی بیش از 96٪ داده ها بین -0.69 تا 7.89 میباشدند

ضریب تغییرات

$$\rho = \frac{\sigma}{\mu} = \frac{2.14}{3.6} = 0.46 = 46\%$$

یعنی پراکندگی داده‌ها معمولی است

همیشه $0 \leq \rho \leq 1$ است

اگر ρ صفر شود یعنی داده ها اصلا پراکنده نیستند داده‌ها بسیار بهم نزدیک هستند

اگر ρ یک شود یعنی داده ها خیلی پراکنده هستند

$$\frac{\sum(x_i - \bar{x}) * f_i}{\sum f_i} = \frac{(1 - 3.6) * 1 + (2 - 3.6) * 2 + (3 - 3.6) * 4 + (5 - 3.6) * 2 + (9 - 3.6) * 1}{1 + 2 + 4 + 2 + 1}$$

$$= \frac{-2.6 - 3.2 - 2.4 + 2.8 + 5.4}{1 + 2 + 4 + 2 + 1} = \frac{0}{10} = 0$$

جدول در بالا تکمیل شد

نسبی r یعنی اینکه مثلاً از داده به مقدار 3 بتعداد 4 تا داریم یعنی 40٪ داده ها 3 هستند
نسبی تجمعی R یعنی اینکه مثلاً از داده به مقدار 3 و کمتر از 3 بتعداد 7 تا داریم یعنی 70٪ داده ها 3 یا کمتر از 3 هستند

15) در یک منطقه میزان مبتلایان به ازدیاد قندخون در چند سال گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال 1399 چقدر میباشد

| | | | | | | | |
|-------|------|------|------|------|------|------|------|
| X=سال | 1394 | 1395 | 1396 | 1397 | 1398 | 1399 | 1400 |
| مصرف | 2 | 2 | 3 | 4 | 5 | ?? | ?? |

سال مبنا را سال 1396 در نظر میگیریم و همه X ها را از 1396 کم میکنیم

| | | | | | | | |
|-------|----|----|---|----|----|----|----|
| X=سال | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| مصرف | 2 | 2 | 3 | 4 | 5 | ?? | ?? |

$$y = a + bx \quad , \quad b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} \quad , \quad \bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$= \frac{(-2 * 2) + (-1 * 2) + (0 * 3) + (1 * 4) + (2 * 5) - \frac{(-2 - 1 + 0 + 1 + 2) * (2 + 2 + 3 + 4 + 5)}{5}}{(-2)^2 + (-1)^2 + 0^2 + (1)^2 + (2)^2 - \frac{(-2 - 1 + 0 + 1 + 2)^2}{5}}$$

$$b = \frac{-4 - 2 + 0 + 4 + 10 - \frac{0 * 16}{5}}{4 + 1 + 0 + 1 + 4 - \frac{0}{5}} \quad b = \frac{8}{10} = 0.8$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-2 - 1 + 0 + 1 + 2}{5} \quad \bar{x} = 0$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{2 + 2 + 3 + 4 + 5}{5} \quad \bar{y} = \frac{16}{5} = 3.2$$

$$\bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x} \rightarrow a = 3.2 - 0.8 * 0 \rightarrow a = 3.2 \rightarrow y = a + bx \rightarrow y = 3.2 + 0.8x$$

$$1399 - 1396 = 3 \rightarrow y = 3.2 + 0.8x \rightarrow y = 3.2 + (0.8 * 3) = 3.2 + 2.4 = 5.6$$

پس برای سال 1399 پیش بینی به مقدار 5.6 میشود

$$1400 - 1396 = 4 \rightarrow y = 3.2 + 0.8x \rightarrow y = 3.2 + (0.8 * 4) = 3.2 + 3.2 = 6.4$$

پس برای سال 1400 پیش بینی به مقدار 6.4 میشود

16) در جدول زیر در خصوص X میانگین 4.7 واریانس 5.7 قبلا خودمان محاسبه کرده ایم اصلا مجدد محاسبه نکنید . در خصوص داد های y که با یک نسبت به x تغییر یافته اند جوابهای میانگین واریانس و انحراف معیار y بدون استفاده از فرمولهای اصلی و سریع از طریق جوابهای x بنویسید

| | | | | | |
|-------------|----|----|----|----|----|
| $x =$ نمره | 2 | 4 | 5 | 6 | 7 |
| $f =$ تعداد | 20 | 10 | 40 | 20 | 10 |

| | | | | | |
|-------------|----|----|----|----|----|
| $y =$ وزن | 62 | 64 | 65 | 66 | 67 |
| $f =$ تعداد | 20 | 10 | 40 | 20 | 10 |

قانونی داریم که اگر داده ها به یک مقدار یکسان اضافه شوند میانگین هم بهمان مقدار اضافه میشود ولی واریانس و انحراف معیار ثابت میماند

در داده های فوق هر یک از داده جدید باندازه 60 واحد بیش از داده قدیم است در خصوص x ها خود مسئله عنوان کرده که

$$\bar{x} = 4.7 \quad \sigma_x = 2.38$$

طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 60$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 60 = 4.7 + 60 \rightarrow \bar{y} = 64.7$$

$$\sigma_y^2 = \sigma_x^2 \rightarrow \sigma_y = 5.7$$

$$\sigma_y = \sigma_x = \sqrt{5.7} = 2.38 \rightarrow \sigma_y = 2.38$$

چون داده های جدید همگی 60 واحد اضافه شده اند پس میانگین داده های جدید 60 واحد بیشتر از میانگین داده های قبلی میشود و انحراف معیار داده های جدید با انحراف معیار داده های قدیم هیچ فرقی نمیکند

17) آمار توصیفی و آمار استنباطی را توضیح دهید

آمار توصیفی، مجموعه ای از روش هایی است که برای سازمان دهی، خلاصه کردن، تهیه جدول، رسم نمودار، توصیف و تفسیر داده های جمع آوری شده از نمونه آماری به کار گرفته می شود - مراحل آن : خلاصه کردن داده ها و توصیف الگوی کلی - فشرده کردن داده ها در قالب جدول های آماری - نمایش آن ها به وسیله نمودار - محاسبه شاخص های آماری

آمار استنباطی مشخص می کند که آیا الگوها و فرآیندهای کشف شده در نمونه، در جامعه آماری هم کاربرد دارد یا خیر. بنابراین، آمار استنباطی راجع به ویژگی ها و پارامترهای مربوط به جامعه آماری تحقیق و کیفیت ارتباط بین مفاهیم و متغیرها می باشد. بدین ترتیب، می توان گفت که از آمار استنباطی در تجزیه و تحلیل مقایسه ای و رابطه ای (علی - همبستگی) استفاده می شود.

18) انواع مقیاسها را نام برده و توضیح ببرید

مقیاس اسمی (NOMINAL) - کد گذاری عددی - اعداد نسبت داده شده ارجحیت و برتری ندارند (مثل جنسیت - زن 2 مرد 1)
 مقیاس رتبه ای (ORDINAL) (ترتیبی) - با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند ولی تناسب ندارند (مثل پرسشنامه لیکرت شامل موافق 3 - بی نظر 2 - مخالف 1)
 مقیاس نسبتی (SCALE) (وزنی) با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند و تناسب هم دارند (مثل سن و مثل وزن - مثلا وزن احمد 35 و وزن عباس 70 کیلوگرم یعنی عباس بیشتر از احمد است و به نسبت دو برابر هم میباشد)

مقیاس فاصله‌ای - مثلاً به یک بیمار بگوییم اگر حداکثر درد مثل یک خط کش عدد 100 باشد درد شما چقدر است و بیمار بگوید 65

19) جمعیت چیست نمونه چیست

جمعیت: مجموعه تمام عناصری که دارای یک یا چند ویژگی مشترک باشند (مثلاً در یک دانشگاه با 400 دانشجو اعداد وزن تمام 400 نفر دانشجو را داشته باشیم)

نمونه: بخشی از جمعیت می‌باشد (مثلاً در یک دانشگاه با 400 دانشجو اعداد وزن 80 نفر دانشجو را داشته باشیم)

20) -نمرات دانشجویان با داده‌های زیر را در جدولی دیگر مرتب کنید. الف) میانۀ داده‌ها چه مقدار است؟ (چرا) ب- مد(نما) داده‌ها چه مقدار است. ج-

صدک 85 داده‌های را معلوم کنید -د- میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید

با تکمیل جدول فراوانی نسبی و فراوانی نسبی تجمعی را محاسبه کنید و با یک مثال نسبی‌ها را توضیح دهید

| | | | | | | |
|-----------|----|----|----|----|----|----|
| X = نمره | 15 | 12 | 20 | 11 | 17 | 14 |
| f = تعداد | 2 | 4 | 1 | 1 | 1 | 1 |

حل: مد داده 14 می‌باشد زیرا بیشترین تکرار (تعداد) را دارد

برای میانۀ و چندک حتماً باید مرتب کنیم و F_i تشکیل دهیم

| | | | | | | |
|-----------|----|----|----|----|----|----|
| X = نمره | 11 | 12 | 14 | 15 | 17 | 20 |
| f = تعداد | 1 | 4 | 1 | 2 | 1 | 1 |
| F_i | 1 | 5 | 6 | 8 | 9 | 10 |

در میانۀ $p=1/2$ است

$$n * p = 10 * \frac{1}{2} = 5 \quad 5+ \rightarrow F_i = 6 \rightarrow X_i = 14 \text{ میانۀ}$$

در صدک 85 خواهیم داشت $p=85/100$

$$n * p = 10 * \frac{85}{100} = 8.5 \quad 8.5+ \rightarrow F_i = 9 \rightarrow X_i = 17$$

یعنی 85 درصد داده‌ها 17 یا کمتر از 17 می‌باشند

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{(11*1)+(12*4)+(14*1)+(15*2)+(17*1)+(20*1)}{1+4+1+2+1+1} = 14 \text{ میانگین}$$

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{(11-14)^2*1+(12-14)^2*4+(14-14)^2*1+(15-14)^2*2+(17-14)^2*1+(20-14)^2*1}{1+4+1+2+1+1} = 7.2 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{7.2} = 2.68 \text{ انحراف معیار}$$

21) - در یک منطقه میزان مصرف سالیانه برق در چند سال گذشته بشرح ذیل میلیون مگاوات ساعت می‌باشد معادله رخط رگرسیون را نوشته پیش

بینی سال 1395 چقدر می‌باشد

| | | | | | | |
|------|------|------|------|------|------|------|
| مصرف | 5 | 6 | 6 | 7 | 8 | ?? |
| | 1390 | 1391 | 1392 | 1393 | 1394 | 1395 |

حل: سال 92 را مبنای صفر قرار می‌دهیم تا راحتتر عملیات ضرب و تقسیم صورت پذیرد

| | | | | | | |
|--------|----|----|---|---|---|----|
| Y=مصرف | 5 | 6 | 6 | 7 | 8 | ?? |
| X=سال | -2 | -1 | 0 | 1 | 2 | 3 |

$$y = a + bx$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}, \quad a = \bar{y} - b\bar{x} \Rightarrow y = a + bx$$

$$y = a + bx, \quad b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}, \quad \bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$= \frac{(5 * -2) + (6 * -1) + (6 * 0) + (7 * 1) + (8 * 2) - \frac{(-2 - 1 + 0 + 1 + 2) * (5 + 6 + 6 + 7 + 8)}{5}}{((-2)^2 + (-1)^2 + 0^2 + (1)^2 + (2)^2 - \frac{(-2 - 1 + 0 + 1 + 2)^2}{5}}$$

$$b = \frac{-10 - 6 + 0 + 7 + 16 - \frac{0 * 32}{5}}{4 + 1 + 0 + 1 + 4 - \frac{0}{5}} \quad b = \frac{7}{10} = 0.7$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-2 - 1 + 0 + 1 + 2}{5} \quad \bar{x} = 0$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{5 + 6 + 6 + 7 + 8}{5} \quad \bar{y} = \frac{32}{5} \quad \bar{y} = 6.4$$

$$a = \bar{y} - b\bar{x} \rightarrow a = 6.4 - 0.7 * 0 \rightarrow a = 6.4$$

$$y = a + bx \rightarrow y = 6.4 + 0.7x \rightarrow x = 3 \rightarrow 1395 \equiv 3 \rightarrow y = 6.4 + (0.7 * 3) = 8.5$$

22) مصرف یک دارو در سه سال گذشته مقادیر زیر بوده در سالهای بعد پیش بینی نمایید

| | | | | | |
|----------|----|----|----|----|----|
| سال = x | 93 | 94 | 95 | 96 | 97 |
| مصرف = y | 1 | 2 | 4 | ? | ? |

حل: ابتدا در جدول مقادیر x قدیمی را Xm نام گذاشته و آنرا را تغییر میدهیم مثلا همه را از 94 کم میکنیم

| | | | |
|----------|----|----|----|
| سال = Xm | 93 | 94 | 95 |
| x | -1 | 0 | +1 |
| مصرف = y | 1 | 2 | 4 |

$$y = a + bx$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i * \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} = \frac{(-1 * 1) + (0 * 2) + (1 * 4) - \frac{(-1 + 0 + 1)(1 + 2 + 4)}{3}}{((-1)^2 + (0)^2 + (1)^2) - \frac{(-1 + 0 + 1)^2}{3}} = \frac{3}{2} = 1.5$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-1 + 0 + 1}{3} = 0 \quad \bar{y} = \frac{\sum y_i}{n} = \frac{1 + 2 + 4}{3} = \frac{7}{3} = 2.33$$

$$\bar{y} = a + b\bar{x} \quad 2.33 = a + (1.5 * 0) \quad a = 2.33$$

$$y = 2.33 + 1.5x$$

$$Xm = 96 \rightarrow x = 96 - 94 = 2 \rightarrow y = 2.33 + (1.5 * 2) = 5.33$$

$$Xm = 97 \rightarrow x = 97 - 94 = 3 \rightarrow y = 2.33 + (1.5 * 3) = 6.83$$

23) در یک منطقه میزان مبتلایان به ازدیاد قندخون در چند سال گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال 1397 چقدر میباشد

| | | | | |
|---------|------|------|------|------|
| سال = x | 1393 | 1394 | 1395 | 1397 |
| مصرف | 5 | 7 | 8 | ?? |

حل: ابتدا در جدول مقادیر x قدیمی را Xm نام گذاشته و آنرا را تغییر میدهیم مثلا همه را از 1394 کم میکنیم

| | | | |
|----------|------|------|------|
| سال = Xm | 1393 | 1394 | 1395 |
| x | -1 | 0 | +1 |
| مصرف = y | 5 | 7 | 8 |

$$y = a + bx$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i * \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} = \frac{(-1 * 5) + (0 * 7) + (1 * 8) - \frac{(-1 + 0 + 1)(5 + 7 + 8)}{3}}{((-1)^2 + (0)^2 + (1)^2) - \frac{(-1 + 0 + 1)^2}{3}} = \frac{3}{2} = 1.5$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-1 + 0 + 1}{3} = 0 \quad \bar{y} = \frac{\sum y_i}{n} = \frac{5 + 7 + 8}{3} = \frac{20}{3} = 6.67$$

$$\bar{y} = a + b\bar{x} \quad 6.67 = a + (1.5 * 0) \quad a = 6.67$$

$$y = 6.67 + 1.5x$$

$$Xm = 1397 \rightarrow x = 1397 - 1394 = 3 \rightarrow y = 6.67 + (1.5 * 3) = 11.17$$

24) یک سکه را دو بار پرتاب میکنیم برای رو آمدن (شیر) جدول تابع چگالی و تابع توزیع را بنویسید و منحنی هر دو تابع را رسم کنید امید شیر بدست آورید

حل: اگر شیر = H و خط = T بنامیم

$$S = \{HH, HT, TH, TT\}$$

تعداد کل حالات چهار تا است و احتمال هر حالت یکی از چهار تا است

| | | | | |
|--------------|------|-----|-----|---------|
| تعداد شیر | X | 0 | 1 | 2 |
| چگالی احتمال | f(x) | 1/4 | 2/4 | 1/4 |
| چگالی توزیع | F(x) | 0 | 1/4 | 3/4 |
| | | | | (4/4)=1 |

$$\text{امید} = E(x) = \sum x * f(x) = \left(0 * \frac{1}{4}\right) + \left(1 * \frac{2}{4}\right) + \left(2 * \frac{1}{4}\right) = 1$$

منحنی هر دو تابع

=====

25) آمار توصیفی و آمار استنباطی را توضیح دهید

آمار توصیفی، مجموعه ای از روش هایی است که برای سازمان دهی، خلاصه کردن، تهیه جدول، رسم نمودار، توصیف و تفسیر داده های جمع آوری شده از نمونه آماری به کار گرفته می شود - مراحل آن : خلاصه کردن داده ها و توصیف الگوی کلی - فشرده کردن داده ها در قالب جدول های آماری - نمایش آن ها به وسیله نمودار - محاسبه شاخص های آماری
آمار استنباطی مشخص می کند که آیا الگوها و فرآیندهای کشف شده در نمونه، در جامعه آماری هم کاربرد دارد یا خیر. بنابراین، آمار استنباطی راجع به ویژگی ها و پارامترهای مربوط به جامعه آماری تحقیق و کیفیت ارتباط بین مفاهیم و متغیرها می باشد. بدین ترتیب، می توان گفت که از آمار استنباطی در تجزیه و تحلیل مقایسه ای و رابطه ای (علی - همبستگی) استفاده می شود.

26) آمارهای توصیفی عمده را نام ببرید

اندازه های گرایش به مرکز (میانگین، میانه، نما)

اندازه های پراکندگی (دامنه تغییرات، واریانس، انحراف استاندارد)

اندازه های وضعیت نسبی (رتبه درصدی، نمره انحراف متوسط)

اندازه های رابطه ای (ضریب همبستگی پیرسون، اسپیرمن)

27) در توزیع نرمال رابطه میانگین با انحراف معیار و دو برابر انحراف معیار و فرمولها را توضیح دهید

28) عدد ضریب همبستگی چه محدوده ای دارد - اگر ضریب همبستگی منفی شود چه معنی دارد - اگر ضریب همبستگی مثبت شود چه معنی دارد

29) تنظیم پرسشنامه آماری سوالات کمکی و سوالات اصلی چیست

30) انواع مقیاسها در آمار توصیفی را نام ببرید هر کدام نیم خط توضیح دهید

مقیاس اسمی (NOMINAL) - کد گذاری عددی - اعداد نسبت داده شده ارجحیت و برتری ندارند (مثل جنسیت - زن 2 مرد 1)

مقیاس رتبه ای (ORDINAL) (ترتیبی) - با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند ولی تناسب ندارند (مثل پرسشنامه لیکرت شامل موافق 3 - بی نظر 2 - مخالف 1)

مقیاس نسبی (SCALE) (وزنی) با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند و تناسب هم دارند (مثل سن و مثل وزن - مثلا وزن احمد

35 و وزن عباس 70 کیلوگرم یعنی عباس بیشتر از احمد است و به نسبت دو برابر هم میباشد)

مقیاس فاصله ای - مثلا به یک بیمار بگوییم اگر حداکثر درد مثل یک خط کش عدد 100 باشد درد شما چقدر است و بیمار بگوید 65

31) انواع روشهای مقیاس گذاری را نام ببرید هر کدام نیم خط توضیح دهید

• انواع مقیاس های مقایسه ای

مقیاس مجموع ثابت: مثلاً ممکن است از پاسخ دهندگان بخواهند 100 امتیاز را به پنج ویژگی اختصاص دهند. نتایج این نوع مقیاس معمولاً به صورت

درصدی عنوان می شود.

آرایش رتبه ای: مثال 10 کتاب به خواننده داده شده و از وی خواسته می شود به کتابی که علاقه بیشتری به آن دارد رتبه یک و به کتابی که اصلاً به آن علاقه

ندارد نمره 10 دهد. این روش برای اندازه‌گیری اولویت کتابها و نیز ویژگی‌های کتاب به‌کار می‌رود. مقایسه جفتی: در این روش یک جفت گزینه به مصرف‌کننده ارائه شده و از او خواسته می‌شود براساس برخی معیارها یکی از آنها را انتخاب کند.

• انواع مقیاس‌های غیرمقیاسه‌ای
طیف لیکرت: معمولاً از 5 قسمت تشکیل شده است (کاملاً موافقم - موافقم - تاحدودی - مخالفم - کاملاً مخالفم) که براساس هدف و روش تحقیق می‌توان کلمات گویه‌ها را عوض نمود
افتراق معنایی: مقیاسی هفت امتیازی و دو قطبی است، مثلاً می‌توان از راضی و ناراضی به‌عنوان نقاط ابتدا و انتها استفاده کرد.
مقیاس استاپل: مقیاس استاپل از معیاری واحد در میانه بازه مقادیر اعداد زوج از 5- تا 5+ و بدون نقطه صفر تشکیل شده است.
مقیاس رتبه‌ای پیوسته: این مقیاس این امکان را برای پاسخ‌دهنده فراهم می‌آورد که به جای اینکه از میان پاسخ‌های از پیش تعیین شده پاسخ مورد نظر را انتخاب کند، در هر نقطه از یک محور علامتی بگذارد.
طیف بوگاردوس: کاربرد این طیف در مواقعی است که مثلاً دو گروه مختلف با دو فرهنگ و زبان متفاوت در کنار هم زندگی می‌کنند (مثلاً در سنجش نگرش به شغل/ طبقه اجتماعی/ گروههای مذهبی) این طیف بکار می‌رود.
32) میانگین - واریانس - انحراف معیار - کوواریانس - ضریب همبستگی توضیح دهید - دو نمونه ضریب همبستگی را نام ببرید و موارد استفاده هر یک را بنویسید

33) چهار مورد از انواع نمودارها نام ببرید و هر یک نیم خط توضیح دهید و شکل نمودار را ترسیم کنید
نمودار میله ای Bar Chart: برای مشاهده فراوانی داده‌ها و مقایسه داده‌ها نسبت به هم
نمودار دایره‌ای Pie Chart: برای مشاهده سهم هر مورد از داده‌ها
نمودار هیستوگرام Histogram: برای مشاهده نحوه توزیع داده‌ها بصورت پیوسته
نمودار خطی Line Chart: برای نشان دادن رابطه بین دوسری اعداد مثلاً خط رگرسیون
نمودار پراکنندگی Scatter Plot: برای نشان دادن پراکنندگی بین دوسری اعداد مثلاً در رگرسیون
34) نام چهار نرم افزار در آمار

نرم افزارهای Statistical Analysis Software - STATA - R - GAMS - Statgraph - Minitab - Amos - SmartPLS - Eviews - SPSS - Lisrel - PASS (NCSS) - SAS - SPSS (Statistical package for social science)

35) یک پروژه در SPSS ارائه دهید
=====

36) در یک کلاس با 40 دانشجو دارای توزیع نرمال و میانگین نمرات 15 و انحراف معیار 4 میباشد یک دانشجو انتخاب میکنیم. الف) احتمال اینکه نمره دانشجو حداکثر 10 شود چقدر است؟ ب) احتمال اینکه نمره دانشجو بیشتر از 10 شود چقدر است (جدول پیوست شده) ج) احتمال اینکه نمره دقیقاً 10 شود؟
حل: توجه شود مسئله ذکر کرده توزیع نرمال است و توجه شود که جدول برای مقادیر کوچکتر یا مساوی و همچنین جدول برای نرمال استاندارد است

$$X \approx N(15, 4^2) \quad p(x \leq 10) = ? \quad p\left(\frac{x - \mu}{\sigma} \leq \frac{10 - 15}{4}\right) = ?$$

بدینترتیب میگوییم جمعیت به نرمال استاندارد تغییر یافت

$$p(z \leq -1.25) = 0.1056 \quad \text{الف) که از جدول نرمال استاندارد}$$

$$p(x > 10) = 1 - p(z \leq 10) = 1 - p\left(\frac{x - \mu}{\sigma} \leq \frac{10 - 15}{4}\right) = 1 - p(z \leq -1.25) =$$

$$1 - 0.1056 = 0.8944 \quad \text{ب)}$$

$$p(x = 10) = p(x \leq 10) - p(x \leq 9) = p\left(\frac{x - \mu}{\sigma} \leq \frac{10 - 15}{4}\right) - p\left(\frac{x - \mu}{\sigma} \leq \frac{9 - 15}{4}\right) =$$

$$= p(z \leq -1.25) - p(z \leq -1.5) = 0.1056 - 0.0668 = 0.0388 \quad \text{ج)}$$

37) سکه همگنی 400 بار پرتاب میکنیم اگر پرتاب‌ها نرمال باشد

اگر فقط جدول توزیع نرمال را در اختیار داشته باشیم حساب کنید احتمال اینکه؟

الف) کمتر یا مساوی 190 شیر ملاحظه شود

ب) کمتر از 195 شیر مشاهده شود؟

ج) بیشتر از 205 پرتاب شیر مشاهده شود؟

د) بیشتر یا مساوی 205 شیر مشاهده شود

ه) دقیقا 200 شیر مشاهده شود

حل: پرتاب سکه تابع توزیع احتمال دو جمله ای است حال احتمال شیر

$$p = \frac{1}{2} = 0.5$$

در تابع توزیع دو جمله ای میانگین برابر است با

$$\mu = np = 400 * 0.5 = 200$$

در تابع توزیع دو جمله ای انحراف معیار برابر است با

$$\sigma = \sqrt{npq} = \sqrt{400 * \frac{1}{2} * (1 - \frac{1}{2})} = \sqrt{100} = 10$$

چون در صورت سوال مطرح شده که پرتاب ها نرمال می باشد پس میتوان گفت که که میتوان این توزیع دو جمله ای را از طریق جدول توزیع نرمال حل نمود

$$p(x \leq 190) = P\left(\frac{x - \mu}{\sigma} \leq \frac{190 - 200}{10}\right) = p(z \leq -1) = 0.1587$$

$$p(x < 195) = p(x \leq 194) = P\left(\frac{x - \mu}{\sigma} \leq \frac{194 - 200}{10}\right) = p(z \leq -0.6) = 0.2743$$

$$p(x > 205) = 1 - p(x \leq 205) = 1 - P\left(\frac{x - \mu}{\sigma} \leq \frac{205 - 200}{10}\right) = 1 - p(z \leq 0.5) = 1 - 0.6915 = 0.3085$$

$$p(x \geq 205) = 1 - p(x \leq 204) = 1 - P\left(\frac{x - \mu}{\sigma} \leq \frac{204 - 200}{10}\right) = 1 - p(z \leq 0.4) = 1 - 0.6554 = 0.3446$$

$$p(x = 200) = p(x \leq 200) - p(x \leq 199) = P\left(\frac{x - \mu}{\sigma} \leq \frac{200 - 200}{10}\right) - P\left(\frac{x - \mu}{\sigma} \leq \frac{199 - 200}{10}\right) = p(z \leq 0) - p(z \leq -0.1) = 0.5 - 0.4602 = 0.0398$$

38) از سابقه تیر اندازی مشخص شده که 80٪ تیرهایش به هدف میخورد فردا قرار است 100 تیر شلیک کند

الف) احتمال اینکه کمتر یا مساوی 75 تیر به هدف بخورد؟

ب) احتمال اینکه بیشتر یا مساوی 85 تیر به هدف بخورد؟

ج) احتمال اینکه دقیقا مساوی 80 تیر به هدف بخورد؟

==== مرتب با بعد از میانگین ====

51) آیا در جدول زیر رابطه ای بین ساعات استفاده از اینترنت با نمره کسب شده توسط دانشجویان وجود دارد

ب - به روش رتبه بندی اسپیرمن

الف - به روش پیرسون

| | | | | | | | | |
|----|----|----|----|----|----|----|----|---------------------|
| 10 | 14 | 15 | 16 | 13 | 14 | 13 | 10 | اعتیاد به اینترنت x |
| 20 | 15 | 12 | 11 | 15 | 15 | 15 | 18 | نمره درس y |

حل: پیرسون

در امتحان اعداد را کامل بنویسید اینجا خلاصه نوشتم

$$\rho(x, y) = \text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{xy} - \bar{x} \bar{y}}{\sigma_x \sigma_y}$$

محاسبات x

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum x_i}{n} = \frac{10 + 14 + 15 + 16 + 13 + 14 + 10 + 13}{8} = 13.13$$

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{(10 - 13.13)^2 + \dots}{8} = 4.11$$

$$\sigma = \sqrt{4.11} = 2.03$$

محاسبات y مثل بالا

محاسبات x*y مثل بالا نهایتا جدول زیر حاصل میشود

| | میانگین | واریانس | انحراف معیار | | | | | | | | |
|----------------|---------|---------|--------------|----|----|----|----|----|-------|------|------|
| ساعت اینترنت x | 10 | 14 | 15 | 16 | 13 | 14 | 10 | 13 | 13.13 | 4.11 | 2.03 |

| | | | | | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|
| نمره y | 20 | 15 | 12 | 11 | 15 | 15 | 18 | 15 | 15.1 | 7.36 | 2.71 |
| x*y | 200 | 210 | 180 | 176 | 195 | 210 | 180 | 195 | 193 | | |

$$\rho(x, y) = \text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{xy} - \bar{x} \bar{y}}{\sigma_x \sigma_y} = \frac{193 - (13.13 * 15.1)}{2.03 * 2.71} = -0.96$$

ضریب همبستگی پیرسون -0.96 شد یعنی دو سری داده فوق وابستگی شدید در خلاف جهت هم دارند هر چه استفاده اینترنت بیشتر شود نمره دانشجو کمتر میشود

ضریب همبستگی اسپیرمن

| | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|----|----|
| x | 10 | 10 | 13 | 13 | 14 | 14 | 15 | 16 |
| rx | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| Rx | 7.5 | 7.5 | 5.5 | 5.5 | 3.5 | 3.5 | 2 | 1 |

داده سری اول
رتبه سری اول
رتبه سری اول نرم شده

| | | | | | | | | |
|----|----|----|-----|-----|-----|-----|----|----|
| Y | 18 | 20 | 15 | 15 | 15 | 15 | 12 | 11 |
| rY | 2 | 1 | 3 | 4 | 5 | 6 | 7 | 8 |
| Ry | 2 | 1 | 4.5 | 4.5 | 4.5 | 4.5 | 7 | 8 |

داده سری دوم
رتبه سری دوم
رتبه سری دوم نرم شده

| | | | | | | | | |
|-----------|-------|-------|---|---|----|----|----|----|
| Rx-Ry | 5.5 | 6.5 | 1 | 1 | -1 | -1 | -5 | -7 |
| (Rx-Ry)^2 | 30.25 | 42.25 | 1 | 1 | 1 | 1 | 25 | 49 |

$$r_s = 1 - \frac{6 \sum_{i=1}^n D^2}{n(n^2 - 1)} = 1 - \frac{6 * \{5.5^2 + 6.5^2 \dots \dots + (-7)^2\}}{8(8^2 - 1)} = 1 - \frac{6 * 150.5}{504} = -0.79$$

ضریب همبستگی اسپیرمن -0.79 شد یعنی دو سری داده وابستگی زیاد در جهت معکوس دارند

52) تعدادی افراد بالای 60 سال فشار خون و میزان حافظه اندازه گرفتیم جداول زیر حاصل شد

| | | | | | | |
|--------------------------|----|----|----|----|----|----|
| X= میزان سن بالای 60 سال | 3 | 4 | 6 | 8 | 10 | 11 |
| Y= فشار خون | 11 | 11 | 13 | 15 | 17 | 18 |

| | | | | | | |
|--------------------------|----|----|---|---|----|----|
| X= میزان سن بالای 60 سال | 3 | 4 | 6 | 8 | 10 | 11 |
| Z= میزان حافظه | 11 | 10 | 8 | 6 | 5 | 4 |

الف) ضریب تغییرات X و ضریب تغییرات Y و ضریب تغییرات Z را بدست آورید اگر این ضریب به صفر نزدیک باشد چه مفهومی دارد و اگر به 100 نزدیک باشد چه معنی دارد و در خصوص جواب بدست آمده ضریب تغییرات در خصوص X و Y و Z توضیح بنویسید

ب) کوواریانس بین دو سری عدد X و Y را بدست آورید و توضیح بنویسید

ج) کوواریانس بین دو سری عدد X و Z را بدست آورید و توضیح بنویسید

د) ضریب همبستگی بین دو سری عدد X و Y را بدست آورید و در خصوص علامت این ضریب و مقدار آن توضیح بنویسید

ه) ضریب همبستگی بین دو سری عدد X و Z را بدست آورید و در خصوص علامت این ضریب و مقدار آن توضیح بنویسید

برای حل از دو سری داده مسئله، جدول سومی بنام حاصلضرب اعداد متناظر دو سری عدد را بدست میآوریم و سپس میانگین و واریانس و انحراف معیار برای سه سری جدا جدا بدست میآوریم

| | | | | | | | |
|--------------|-----|----|----|----|-----|-----|-----|
| داده سری اول | x | 3 | 4 | 6 | 8 | 10 | 11 |
| داده سری دوم | y | 11 | 11 | 13 | 15 | 17 | 18 |
| | x*y | 33 | 44 | 78 | 120 | 170 | 198 |

$$\bar{x} = \frac{\sum x_i}{n} = \frac{3 + 4 + 6 + 8 + 10 + 11}{6} = 7$$

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} \quad \mu = \bar{x} = \frac{\sum x_i}{n}$$

$$\sigma_x^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{(3-7)^2 + (4-7)^2 + (6-7)^2 + (8-7)^2 + (10-7)^2 + (11-7)^2}{6} = 8.67$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{8.67} = 2.94$$

$$\rho = \frac{\sigma}{\mu} \times 100 = \frac{2.94}{7} * 100 = 42\%$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{11 + 11 + 13 + 15 + 17 + 18}{6} = 14.17$$

$$\sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n} = \frac{(11-14.17)^2 + (11-14.17)^2 + (13-14.17)^2 + (15-14.17)^2 + (17-14.17)^2 + (18-14.17)^2}{6} = 7.47$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{7.47} = 2.73$$

$$\rho = \frac{\sigma}{\mu} \times 100 = \frac{2.73}{14.17} * 100 = 19\%$$

$$\bar{x} \cdot \bar{y} = \frac{\sum x_i y_i}{n} = \frac{33+44+78+120+170+198}{6} = 107.17$$

$$Cov(x, y) = \bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y} = 107.17 - (7 * 14.17) = 8$$

$$\rho = \frac{Cov(x, y)}{\sigma_x \cdot \sigma_y} = \frac{\bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y} = \frac{107.17 - (7 * 14.17)}{2.94 * 2.73} = 0.99$$

بین X و Y همبستگی زیاد در جهت موافق یکدیگر وجود دارد

دو سری داده بعدی مثل روش فوق حل میکنیم

| | | | | | | | |
|----------------|---|----|----|---|---|----|----|
| داده سری سوم | x | 3 | 4 | 6 | 8 | 10 | 11 |
| داده سری چهارم | z | 11 | 10 | 8 | 6 | 5 | 4 |

| | | | | | | |
|-------------|----|----|----|----|----|----|
| $x \cdot z$ | 33 | 40 | 48 | 48 | 50 | 44 |
|-------------|----|----|----|----|----|----|

$$\bar{x} = \frac{\sum x_i}{n} = 7 \quad \sigma^2 = 8.67 \quad \sigma = 2.94 \quad \rho = 42\% \text{ ضریب تغییرات } x$$

$$\bar{z} = \frac{\sum z_i}{n} = \frac{11 + 10 + 8 + 6 + 5 + 4}{6} = 7.33$$

$$\sigma_z^2 = \frac{\sum (z_i - \bar{z})^2}{n} = \frac{(11 - 7.33)^2 + (10 - 7.33)^2 + (8 - 7.33)^2 + (6 - 7.33)^2 + (5 - 7.33)^2 + (4 - 7.33)^2}{6} = 6.56$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{6.56} = 2.56$$

$$\rho = \frac{\sigma}{\mu} \times 100 = \frac{2.56}{7.33} * 100 = 34.9\%$$

$$\overline{x \cdot z} = \frac{\sum x_i z_i}{n} = \frac{33 + 40 + 48 + 48 + 50 + 44}{6} = 43.83$$

$$Cov(x, z) = \overline{x \cdot z} - \bar{x} \cdot \bar{z} = 43.83 - (7 * 7.33) = -7.5$$

$$\rho = \frac{Cov(x, z)}{\sigma_x \cdot \sigma_z} = \frac{\overline{x \cdot z} - \bar{x} \cdot \bar{z}}{\sigma_x \cdot \sigma_z} = \frac{43.83 - (7 * 7.33)}{2.94 * 2.56} = -1$$

بین X و Z همبستگی زیاد در جهت معکوس یکدیگر وجود دارد

53) رابطه بین ساعات استفاده از شبکه اجتماعی با نتیجه آزمایش چربی خون بشرح زیر است

| X | 8 | 7 | 6 | 5 | 4 | 3 | 2 | اعتیاد به شبکه اجتماعی |
|---|-----|-----|-----|-----|-----|-----|-----|------------------------|
| Y | 600 | 530 | 460 | 400 | 350 | 320 | 300 | آزمایش |

کوواریانس (همابستگی) بین دو متغیر فوق را محاسبه کنید و جواب بدست آمده را تحلیل کنید

ضریب همبستگی پیرسون بین دو متغیر فوق را بدست آورده و تحلیل نمایید؟

با توجه به روند نمرات اعتیاد فوق اگر عدد اعتیاد 9 شود عدد آزمایش او را پیش بینی نمایید (رگرسیون)؟

(نمونه حل در جزوه)

54) بین 8 نفر نمرات درس آمار در دانشگاه و نمره ریاضی آنان بشرح زیر است

کوواریانس (همابستگی) بین دو متغیر فوق را محاسبه کنید و جواب بدست آمده را تحلیل کنید

ضریب همبستگی اسپیرمن محاسبه کنید ؟

با توجه به روند نمرات درس ریاضی اگر نمره ریاضی دانشجویی 16 شود نمره درس آمار او را پیش بینی نمایید (رگرسیون)؟

(نمونه حل در جزوه)

| نمره ریاضی دانشگاه | نمره درس آمار در دانشگاه |
|--------------------|--------------------------|
| 13 | 12 |
| 9 | 8 |
| 14 | 13 |
| 14 | 14 |
| 14 | 15 |
| 11 | 11 |
| 11 | 9 |
| 11 | 10 |