

**امتحانات در ۵ گروه برگزار شده است که اعداد اکثر سوالات در هر گروه متفاوت است بعنوان نمونه یک مورد در اینجا کامل حل شده است**

(۱) نام دو نرم افزار در آمار که در این ترم تحصیلی آموزش دیدید را ذکر کنید (نمره)

SPSS (Statistical package for social science)

Excel

Minitab Eviews Statistica Lisrel

Expert choice SAS (Statistical Analysis Software)

NCSS Microfit GAMS

STATA SmartPLS Amos

Statgraph PASS(NCSS)

R

Mathlab

(۲) جمعیت چیست نمونه چیست (نمره)

جمعیت: مجموعه تمام عناصری که دارای یک یا چند ویژگی مشترک باشند  
نمونه: بخشی از جمعیت میباشد

(۳) تعریف آمار چیست (نمره)

روش و چگونگی جمع آوری اطلاعات و بیان آنها در قالب عدد

(۴) آمار توصیفی چیست آمار استنباطی چیست (نمره)

در آمار توصیفی از ویژگی های نمونه آماری تحقیق به همراه شاخص های گرایش به مرکز و یا شاخص های گرایش به پراکندگی می باشد. در حالی که در آمار استنباطی و یا تحلیلی می توان نتایج و یافته های به دست آمده از نمونه آماری را به کل جامعه آماری تحقیق تعمیم داد. به عبارتی، مفهوم کانونی آمار استنباطی، تعمیم پذیری است.

آمار توصیفی زیستی، 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. نقش آمار توصیفی در فرآیند تحلیل آماری بسیار مهم و حیاتی است. آمار توصیفی با خلاصه کردن داده ها، ویژگی های مهم آن ها را نمایان می سازد

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.

(۵) انواع روشهای مقیاس گذاری (measure) را نام ببرید هر کدام نیم خط توضیح دهید (نمره)

مقیاس اسمی (NOMINAL) - کد گذاری عددی - اعداد نسبت داده شده ارجحیت و برتری ندارند (مثل جنسیت - زن ۲ مرد ۱)

مقیاس رتبه ای (ORDINAL) (ترتیبی) - با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند ولی تناسب ندارند (مثل پرسشنامه لیکرت شامل موافق ۳ - بی نظر ۲ - مخالف ۱)

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 ۳۵ و وزن عباس ۷۰ 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 ۱۰۰ باشد درد شما چقدر است و بیمار بگوید ۶۵

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

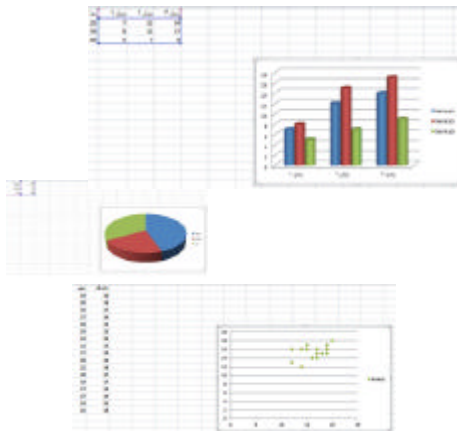
نمودار میله ای Bar Chart: برای مشاهده فراوانی داده ها و مقایسه داده ها نسبت به هم

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



(۷) در محاسبه واریانس اگر توان دوم را فراموش کنیم نتیجه چه خواهد شد (نمره)

صفر

۸) اگر در رگرسیون خطی اعداد  $y$  از چپ به راست از اولین عدد تا جدیدترین عدد  $y$  آرام آرام کم شوند در معادله خط  $y=a+bx$  علامت ضریب خط یعنی  $b$  مثبت میشود یا منفی (۱نمره)

منفی

۹) در محاسبه واریانس یا انحراف معیار ، مقدار جواب بیانگر معدل اعداد است یا پراکندگی اعداد (۱نمره)  
پراکندگی

۱۰) اگر بجای ارائه ده هزار داده عددی بخواهیم شاخصی ارائه دهیم بهترین شاخص (مود) یا (میانه) یا (میانگین با انحراف معیار) کدامیک است (۱نمره)

میانگین با انحراف معیار

۱۱) اگر بخواهیم بجای عدد ۳ و بجای آبی عدد ۲ و بجای قرمز عدد ۵ را بگذاریم این مقیاس گذاری ordinal یا nominal یا scale کدامیک است (۱نمره)

Nominal اسمی

حل ریاضی یک نمونه از سوالات در صفحات بعد

توجه شود این اعداد ممکن است با اعداد سوالات امتحان شما متفاوت باشد

۱۲) نمرات دانشجویان بشرح زیر است. الف) مد(نما) داده ها چه مقدار است (چرا) ب). میانه داده ها چه مقدار است؟ ج) صدک ۱۳ را محاسبه کنید. د) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید ه) جدول فراوانی و فراوانی تجمعی و فراوانی نسبی و تجمعی را تکمیل کنید و در این جدول در خصوص همه اعداد زیر داده ۱۴ یک به یک توضیح دهید (۴۰نمره)

نمرات: ۱۲ و ۱۰ و ۱۴ و ۱۷ و ۱۲ و ۱۵ و ۱۴ و ۱۴ و ۱۷

نمره = x	10	12	14	15	17
تعداد = f	1	2	4	1	2
F	1	3	7	8	10
$r_i$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{1}{10}$	$\frac{2}{10}$
$R_i$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{7}{10}$	$\frac{8}{10}$	$\frac{10}{10}$

مد داده ای بمقدار 14 میباشد زیرا بیشترین تعداد را دارد

$$M=14$$

میانه داده ها

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 4 + 1 + 2) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=7 \rightarrow m=14$$

یعنی نصف داده ها 14 یا کمتر از 14 هستند

صدک ۱۳ داده ها

$$Q \sum f_i = \frac{13}{100}(1 + 2 + 4 + 1 + 2) = \frac{130}{100} = 1.3 \rightarrow \text{در سطر } F_i \rightarrow 1.3+ \rightarrow F_i=3 \rightarrow x=12$$

یعنی 13% داده ها 12 یا کمتر از 12 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(10 * 1) + (12 * 2) + (14 * 4) + (15 * 1) + (17 * 2)}{1 + 2 + 4 + 1 + 2} = \frac{139}{10} = 13.9$$

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

$$\sigma^2 = \frac{(10 - 13.9)^2 * 1 + (12 - 13.9)^2 * 2 + (14 - 13.9)^2 * 4 + (15 - 13.9)^2 * 1 + (17 - 13.9)^2 * 2}{1 + 2 + 4 + 1 + 2} = 4.29 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.29} = 2.07 \text{ انحراف معیار}$$

۴ تا از داده‌ها ۱۴ میباشند

۷ تا از داده‌ها ۱۴ یا کمتر از ۱۴ میباشند

$\frac{4}{10}$  یعنی ۴۰٪ از داده‌ها ۱۴ میباشند

$\frac{7}{10}$  یعنی ۷۰٪ از داده‌ها ۱۴ یا کمتر از ۱۴ میباشند

۱۳) جدول سوال زیر قبلا حل شده است و میانگین 13.2 و واریانس 2.26 میباشد. مجدد حل نکنید

$X_i$	10	12	13	14	16
$f_i$	2	3	6	7	2

الف) انحراف معیار را حساب کنید (ب) بدون حل کردن اگر هنگام درج فرمول واریانس توان دوم را فراموش کنیم نتیجه کسر حاصل چه خواهد شد (محاسبه نکنید فقط جواب بنویسید) 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 و ضریب تغییرات این داده جدید را بدست آورید (بدون استفاده از فرمول اصلی) (۲۰)

$X_i$	310	312	313	314	316
$f_i$	2	3	6	7	2

$$\bar{x} = 13.2 \quad \sigma_x = \sqrt{2.26} = 1.50$$

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

$$\mu \pm \sigma = 13.2 \pm 1.5 = 11.7 \quad \text{تا} \quad 14.7$$

یعنی بیش از ۶۸٪ داده ها بین 11.7 تا 14.7 میباشدند

$$\mu \pm 2\sigma = 13.2 \pm (2 * 1.5) = 13.2 \pm 3 = 10.2 \quad \text{تا} \quad 16.2$$

یعنی بیش از ۹۶٪ داده ها بین 10.2 تا 16.2 میباشدند  
طبق جدول فوق در خصوص  $y$  ها میتوان نوشت که

$$y_i = x_i + 300$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 100 = 13.2 + 300 \rightarrow \bar{y} = 313.2$$

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

$$\sigma_y = \sigma_x = \sqrt{2.26} = 1.5 \rightarrow \sigma_y = 1.5$$

یعنی چون داده های جدید همگی 300 واحد اضافه شده اند پس میانگین داده های جدید 300 واحد بیشتر از میانگین داده های قبلی میشود و انحراف معیار داده های جدید با انحراف معیار داده های قدیم هیچ فرقی نمیکند

ضریب تغییرات

$$\rho = \frac{\sigma}{\mu} = \frac{1.5}{313.2} = 0.0048 = 0.48\%$$

الان  $\rho = 0.48$  عدد خیلی کوچکی است یعنی پراکندگی داده ها کم است

۱۴) در یک منطقه عدد آزمایش خون افراد در چند سال گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال ۱۴۰۲ چقدر میباشد (۲۰نمره)

سال=X	۱۳۹۷	۱۳۹۸	۱۳۹۹	۱۴۰۰	۱۴۰۱	۱۴۰۲
خون	۲	۴	۴	۵	۵	؟؟

همه اعدادی که به نام X داریم ۱۳۹۹ تا کم میکنیم

سال=y	-2	-1	0	+1	+2	+3
خون	۲	۴	۴	۵	۵	؟؟

$$y = a + bx$$

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

$$= \frac{(-2 * 2) + (-1 * 4) + (0 * 4) + (1 * 5) + (2 * 5) - \frac{(-2 - 1 + 0 + 1 + 2)(2 + 4 + 4 + 5 + 5)}{5}}{((-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2) - \frac{(-2 - 1 + 0 + 1 + 2)^2}{5}} = \frac{7}{10}$$

$$= 0.7$$

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

$$\bar{y} = \frac{\sum y_i}{n} = \frac{2 + 4 + 4 + 5 + 5}{5} = \frac{20}{5} = 4$$

$$\bar{y} = a + b\bar{x} \quad 4 = a + (0.7 * 0) \quad a = 4$$

$$y = 4 + 0.7x$$

$$Xm = 1399 \rightarrow y = 1402 - 1399 = 3 \rightarrow y = 4 + (0.7 * 3) = 6.1$$