

سوالات و حل امتحان میانترم 23-05-1400

1) نام دو نرم افزار که در آمار از آن استفاده کردید بنویسید (کنمره)

Excel SPSS (Statistical package for social science)
 SAS (Statistical Analysis Software) Minitab Eviews Statistica Lisrel Expert choice
 NCSS Microfit GAMS STATA SmartPLS Amos R
 Mathlab PASS(NCSS) Statgraph

2) آمار توصیفی و آمار استنباطی را توضیح دهید (کنمره)

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. درحالی که در آمار استنباطی و یا تحلیلی می توان نتایج 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.

3) انواع روشهای مقیاس گذاری را نام ببرید هر کدام نیم خط توضیح دهید (کنمره)

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 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 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

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

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



5) جدول سوال زیر قبلا حل شده است و میانگین 14.1 و واریانس 3.98 میباشد. مجدد حل نکنید

X_i	10	13	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 و ضربت تغییرات این داده جدید را بدست آورید (20)

X_i	110	113	114	115	117
f_i	1	3	2	2	2

$$\bar{x} = 14.1 \quad \sigma_x = \sqrt{3.98} = 1.995$$

P(1)

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

$$\mu \pm \sigma = 14.1 \pm 1.995 = 12.105 \quad \text{تا} \quad 16.095$$

یعنی بیش از 68٪ داده ها بین 12.105 تا 16.095 میباشند

$$\mu \pm 2\sigma = 14.1 \pm (2 * 1.995) = 14.1 \pm 3.99 = 10.11 \quad \text{تا} \quad 18.09$$

یعنی بیش از 96٪ داده ها بین 10.11 تا 18.09 میباشند
طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 100$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 100 = 14.1 + 100 \rightarrow \bar{y} = 114.1$$

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

$$\sigma_y = \sigma_x = \sqrt{3.98} = 1.995 \rightarrow \sigma_y = 1.995$$

چون داده های جدید همگی 100 واحد اضافه شده اند پس میانگین داده های جدید 100 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

$$\rho = \frac{\sigma}{\mu} = \frac{1.995}{114.1} = 0.0175 = 1.7\%$$

الان $\rho = 0.0175$ عدد خیلی کوچکی است یعنی پراکندگی داده ها کم است
همیشه $0 \leq \rho \leq 1$ است. اگر ρ صفر شود یعنی داده ها اصلا پراکنده نیستند دادها بسیار بهم نزدیک هستند
اگر ρ یک شود یعنی داده ها خیلی پراکنده هستند

(6) در یک منطقه میزان مبتلایان به ازدیاد قندخون در چند سال گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال 1401 را محاسبه کنید (20نمره)

سال=x	1395	1396	1397	1398	1399	1401
میزان قند	2	3	4	4	5	??

سال=x	1395	1396	1397	1398	1399	1401
میزان قند	2	3	4	4	5	??
Xn	-2	-1	0	1	2	4

$$y = a + bx$$

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

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

$$= \frac{7}{10} = 0.7$$

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

$$\bar{y} = \frac{\sum y_i}{n} = \frac{2 + 3 + 4 + 4 + 5}{5} = \frac{18}{5} = 3.6$$

P(2)

$$\bar{y} = a + b\bar{x} \quad 3.6 = a + (0.7 * 0) \quad a = 3.6$$

$$y = 3.6 + 0.7x$$

$$Xm = 1400 \rightarrow x = 1401 - 1397 = 4 \rightarrow y = 3.6 + (0.7 * 4) = 6.4$$

7) نمرات دانشجویان بشرح زیر است داده‌ها را در جدول زیر مرتب کنید. الف) مد(نما) داده‌ها چه مقدار است (چرا) ب). میانه داده‌ها چه مقدار است؟ - ج) صدک چهارم داده‌های را معلوم کنید - د) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید . و) با تکمیل جدول، فراوانی نسبی و فراوانی نسبی تجمعی را محاسبه کنید و با یک مثال نسبی‌ها را توضیح دهید .. نمرات : 15 و 15 و 12 و 11 و 19 و 17 و 13 و 15 و 10 و 13 (20نمره)

نمره = x							
تعداد = f							

حل:

x داده	10	11	12	13	15	17	19
f فراوانی	1	1	1	2	3	1	1
F فراوانی تجمعی	1	2	3	5	8	9	10
f _i فراوانی نسبی $I = \frac{f_i}{\sum f_i}$	1/10=10%	1/10=10%	1/10=10%	2/10=20%	3/10=30%	1/10=10%	1/10=10%
R فراوانی نسبی تجمعی $R = \frac{\sum f_{i_i}}{\sum f_i}$	1/10=10%	2/10=20%	3/10=30%	5/10=50%	8/10=80%	9/10=90%	10/10=100%

مد داده ای بمقدار 15 میباشد زیرا بیشترین تعداد را دارد

$$M=15$$

میانه داده‌ها

$$Q \sum f_i = \frac{1}{2}(1 + 1 + 1 + 2 + 3 + 1 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=8 \rightarrow m=15$$

یعنی نصف داده‌ها 15 یا کمتر از 15 هستند

صدک 14 داده‌ها

$$Q \sum f_i = \frac{14}{100}(1 + 1 + 1 + 2 + 3 + 1 + 1) = \frac{14}{100} = 1.4 \rightarrow \text{در سطر } F_i \rightarrow 1.4+ \rightarrow F_i=2 \rightarrow x=11$$

یعنی 14% داده‌ها 11 یا کمتر از 11 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i}$$

$$= \frac{(10 * 1) + (11 * 1) + (12 * 1) + (13 * 2) + (15 * 3) + (17 * 1) + (19 * 1)}{1 + 1 + 1 + 2 + 3 + 1 + 1}$$

$$= \frac{140}{10} = 14$$

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

P(3)

$$\sigma^2 = \frac{(10-14)^2 * 1 + (11-14)^2 * 1 + (12-14)^2 * 1 + (13-14)^2 * 2 + (15-14)^2 * 3 + (17-14)^2 * 1 + (19-14)^2 * 1}{1+1+1+2+3+1+1}$$

واریانس 6.8 =

$$\sigma = \sqrt{\sigma^2} = \sqrt{6.8} = 2.61 \quad \text{انحراف معیار}$$

جدول در بالا تکمیل شد

نسبی r یعنی اینکه مثلاً از داده به مقدار 13 بتعداد 2 تا داریم یعنی 20٪ داده ها 13 هستند
 نسبی تجمعی R یعنی اینکه مثلاً از داده به مقدار 13 و کمتر از 13 بتعداد 5 تا داریم یعنی 50٪ داده ها 13 یا کمتر از 13 هستند

8) یک دانشجو با 80 واحد را با توزیع نرمال و میانگین نمرات 15 و واریانس 4 گذرانده است، این دانشجو در یک درس در ترم بعدی الف) احتمال اینکه نمره دانشجو حداکثر 14 شود چقدر است؟ ب) احتمال اینکه نمره دانشجو بیشتر از 17 شود چقدر است ج) احتمال اینکه نمره دقیقاً 16 شود؟ (جدول بشرح زیر در نظر بگیرید 20 نمره)

$$p(z \leq 2) = 0.98, \quad p(z \leq 1.5) = 0.93, \\ p(z \leq 1) = 0.85, \quad p(z \leq 0.5) = 0.7, \quad p(z \leq 0) = 0.5 \\ p(z \leq -2) = 0.02, \quad p(z \leq -1.5) = 0.07, \quad p(z \leq -1) = 0.15, \quad p(z \leq -0.5) = 0.3$$

$$P(x \leq 14) = P\left(\frac{x - \mu}{\sigma} \leq \frac{14 - 15}{2}\right) = P(z \leq -0.5) = 0.3$$

$$P(x > 18) = 1 - P(x \leq 18) = 1 - P\left(\frac{x - \mu}{\sigma} \leq \frac{18 - 15}{2}\right) = 1 - P(z \leq 1.5) = 1 - 0.97 = 0.03$$

$$P(x = 16) = P(x \leq 16) - P(x \leq 15) = P\left(\frac{x - \mu}{\sigma} \leq \frac{16 - 15}{2}\right) - P\left(\frac{x - \mu}{\sigma} \leq \frac{15 - 15}{2}\right) \\ = P(z \leq 0.5) - P(z \leq 0) = 0.7 - 0.5 = 0.2$$