

آمار زیستی - سوال و پاسخ میانترم آذر 1400

سوال 5	انحراف معیار	$\mu \pm \sigma$	$\mu \pm 2\sigma$	میانگین	انحراف معیار	ضریب تغییرات
پاسخ	1.939	12.261	10.322	114.2	1.939	0.017=1.7%
		16.139	18.078			

سوال 6	مود	میانه	صدک 23	میانگین	واریانس	انحراف معیار
پاسخ	15	15	11	13.4	5.44	2.33

سوال 7	1401
پاسخ	b = 0.4 3.2

سوال 8	الف	ب	ج
پاسخ	0.3085	0.4013	0.0987

چهار سوال اول هر کدام حد اکثر حدود چهار خط توضیح دهید

(1) نام دو نرم افزار که در آمار از آن استفاده کردید بنویسید (5نمره)

Excel SPSS (Statistical package for social science)

SAS (Statistical Analysis Software) Minitab Eviews Statistica Lisrel Expert choice
 NCSS Microfit GAMS STATA SmartPLS Amos
 Matlab R PASS(NCSS) Statgraph

(2) آمار توصیفی و آمار استنباطی را توضیح دهید (5نمره)

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 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.

(3) انواع روشهای مقیاس گذاری را نام ببرید هر کدام نیم خط توضیح دهید (5نمره)

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

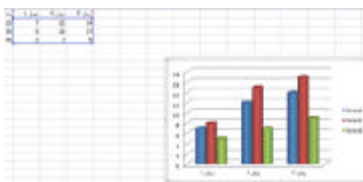
مقیاس رتبه ای (ORDINAL) (ترتیبی) - با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند ولی تناسب ندارند (مثل پرسشنامه لیکرت شامل موافق 3 - بی نظر 2 - مخالف 1)

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

(4) چهار مورد از انواع نمودارها نام ببرید و هر مورد، نیم خط توضیح دهید و شکل نمودار را ترسیم کنید (5نمره)

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 مثلاً

خط رگرسیون

نمودار پراکندگی Scatter Plot: برای نشان دادن پراکندگی بین دوسری

اعداد مثلاً در رگرسیون

5) جدول سوال زیر قبلاً حل شده است و میانگین 14.2 و واریانس 3.76 میباشد. مجدد حل نکنید

x_i	10	13	14	15	17
f_i	1	2	3	2	2

الف) انحراف معیار را حساب کنید

ب) 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. (ج) 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) ضریب تغییرات این داده جدید را بدست آورید (10) (20 نمره)

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

$$\bar{x} = 14.2 \quad \sigma_x = \sqrt{3.76} = 1.939$$

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

$$\mu \pm \sigma = 14.2 \pm 1.939 = 12.261 \quad \text{تا} \quad 16.139$$

یعنی بیش از 68% داده ها بین 12.261 تا 16.139 میباشدند

$$\mu \pm 2\sigma = 14.2 \pm (2 * 1.939) = 14.2 \pm 3.878 = 10.322 \quad \text{تا} \quad 18.078$$

یعنی بیش از 96% داده ها بین 10.322 تا 18.078 میباشدند
طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 100$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 100 = 14.2 + 100 \rightarrow \bar{y} = 114.2$$

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

$$\sigma_y = \sigma_x = \sqrt{3.76} = 1.939 \rightarrow \sigma_y = 1.939$$

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

$$\rho = \frac{\sigma}{\mu} = \frac{1.939}{114.2} = 0.017 = 1.7\%$$

الآن $\rho = 0.017$ عدد خیلی کوچکی است یعنی پراکندگی داده ها کم است

همیشه $0 \leq \rho \leq 1$ است. اگر ρ صفر شود یعنی داده ها اصلاً پراکنده نیستند داده ها بسیار بهم نزدیک هستند

اگر ρ یک شود یعنی داده ها خیلی پراکنده هستند

6) داده‌های زیر را داریم . الف) مد(نما) داده‌ها محاسبه کنید ب- میانه داده‌ها محاسبه کنید ج- صدک 23 داده‌ها را محاسبه کنید د- میانگین و واریانس و انحراف معیار داده‌ها را محاسبه کنید . ه) در ادامه در جدولی مرتب شده ، فراوانی تجمعی ، فراوانی نسبی و فراوانی نسبی تجمعی را بنویسید . (40نمره)

x = نمره	13	11	17	9	15
f = تعداد	2	2	1	1	4

x داده	9	11	13	15	17
f فراوانی	1	2	2	4	1
F فراوانی تجمعی	1	3	5	9	10
r = $\frac{f_i}{\sum f_i}$ فراوانی نسبی	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{1}{10}$
R = $\frac{\sum f_{i0.a}}{\sum f_i}$ فراوانی نسبی تجمعی	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{5}{10}$	$\frac{9}{10}$	$\frac{10}{10}$

مد یا نما : همان داده ای است که بیشترین فراوانی(تعداد) را دارد که میشود 15

$$M=15$$

میانه داده‌ها

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 2 + 4 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=9 \rightarrow m=15$$

صدک هفتاد دوم

$$Q \sum f_i = \frac{23}{100}(1 + 2 + 2 + 4 + 1) = \frac{23}{100} = 2.3 \rightarrow \text{در سطر } F_i \rightarrow 2.3+ \rightarrow F_i=3 \rightarrow m=11$$

یعنی 23% داده‌ها 11 یا کمتر از 11 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(9 * 1) + (11 * 2) + (13 * 2) + (15 * 4) + (17 * 1)}{1 + 2 + 2 + 4 + 1} = \frac{134}{10} = 13.4$$

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

$$\sigma^2 = \frac{(9 - 13.4)^2 * 1 + (11 - 13.4)^2 * 2 + (13 - 13.4)^2 * 2 + (15 - 13.4)^2 * 4 + (17 - 13.4)^2 * 1}{1 + 2 + 2 + 4 + 1} = \frac{54.4}{10} = 5.44 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{5.44} = 2.33 \text{ انحراف معیار}$$

$$\mu \pm \sigma = 13.4 - 2.33 = 11.7, \quad 13.4 + 2.33 = 15.73$$

68% داده‌ها بین 11.7 تا 15.73 میباشد

$$\mu \pm 2\sigma = 13.4 - (2 * 2.33) = 8.74, \quad 13.4 + (2 * 2.33) = 18.06$$

96% داده‌ها بین 8.74 تا 18.06 میباشد

نسبی r یعنی اینکه مثلا از داده به مقدار 13 بتعداد 2 تا داریم یعنی 20% داده‌ها 13 هستند

نسبی تجمعی R یعنی اینکه مثلا از داده به مقدار 13 و کمتر از 13 بتعداد 5 تا داریم یعنی 50% داده‌ها 13 یا کمتر از 13 هستند

$$\rho = \frac{\sigma}{\mu} = \frac{2.33}{13.4} = 0.1739 = 17.39\% \text{ ضریب تغییرات}$$

الان $\rho = 0.173$ عدد کوچکی است یعنی پراکندگی داده‌ها نسبتا کم است - همیشه $0 \leq \rho \leq 1$ است . اگر ρ صفر شود یعنی داده‌ها اصلا پراکنده

نیستند داده‌ها بسیار بهم نزدیک هستند - اگر ρ یک شود یعنی داده‌ها خیلی پراکنده هستند

7) در یک منطقه میزان مصرف برق در سالهای گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال

بعد چقدر میباشد (15نمره) (صفر نمره)

X	1396	1397	1398	1399	1400	1401
Y	1	2	2	2	3	?

X	1396	1397	1398	1399	1400	1401
Y	1	2	2	2	3	?
	-2	-1	0	1	2	

$$y = a + bx$$

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

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

$$= \frac{4}{10} = 0.4$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-2 - 1 + 0 + 1 + 2}{5} = 0 \quad \bar{y} = \frac{\sum y_i}{n} = \frac{1 + 2 + 2 + 2 + 3}{5} = \frac{10}{5} = 2$$

$$\bar{y} = a + b\bar{x} \quad 2 = a + (0.4 * 0) \quad a = 2$$

$$y = 2 + 0.4x$$

$$X_m = 1401 \rightarrow x = 1401 - 1398 = 3 \rightarrow y = 2 + (0.4 * 3) = 3.2$$

8) یک دانشجو 80 واحد را با توزیع نرمال و میانگین نمرات 17 و واریانس 16 گذرانده است، این دانشجو در یک درس

در ترم بعدی. الف) احتمال اینکه نمره دانشجو حداکثر 15 شود چقدر است؟ ب) احتمال اینکه نمره دانشجو بیشتر از

18 شود چقدر است ج) احتمال اینکه نمره دقیقا 17 شود؟ (جدول از جزوه استفاده کنید) (15نمره) (20 نمره)

$$p(x \leq 15) = P\left(\frac{x - \mu}{\sigma} \leq \frac{15 - 17}{4}\right) = p(z \leq -0.5) = 0.3085$$

$$p(x > 18) = 1 - P(x \leq 18) = 1 - \left(\frac{x - \mu}{\sigma} \leq \frac{18 - 17}{4}\right) = 1 - p(z \leq 0.25) = 1 - 0.5987 = 0.4013$$

$$p(x = 17) = p(x \leq 17) - p(x \leq 16) = P\left(\frac{x - \mu}{\sigma} \leq \frac{17 - 17}{4}\right) - P\left(\frac{x - \mu}{\sigma} \leq \frac{16 - 17}{4}\right) = p(z \leq 0) - p(z \leq -0.25)$$

$$= 0.5 - 0.4013 = 0.0987$$