

امتحان پایانترم کارگاه آمار زیستی 7 سوال (ارزش امتحان پایانترم 3 نمره از 20 نمره میباشد) - حتما در یک محل خالی در وسط هر برگ پاسخ نام و فامیل نوشته و همه پاسخها در یک فایل پی دی اف ارسال کنید

1) نرم افزار اکسل و SPSS کدامیک از شاخصهای آماری زیر را انجام میدهند(10نمره)

مود - میانه - میانگین - واریانس - انحراف معیار - رگرسیون

تمام شاخصها

2) دانشجویی به اشتباه در یک سری داده آماری، در فرمول واریانس، توان 2 را در محاسبات فراموش نموده جواب چه خواهد شد(10نمره)

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

صفر

3) انواع روشهای مقیاس گذاری را نام ببرید هر کدام نیم خط توضیح دهید(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 لیکرت شامل موافق 3 - بی

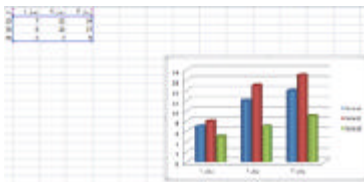
نظر 2 - مخالف 1)

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 فاصله‌ای - 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(10نمره)



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

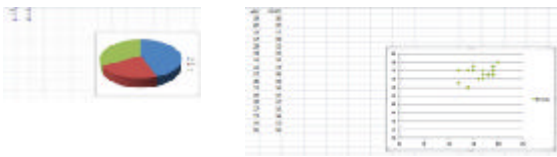
نمودار دایره ای Pie Chart: برای مشاهده سهم هر مورد از داده ها

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



! در حل سه سوال زیر باید است حداقل دو رقم اعشار رعایت شود!

5) جدول سوال زیر قبلا حل شده است و میانگین 13.2 و واریانس 2.36 میباشد. مجدد حل نکنید

X_i	11	13	12	14	17
f_i	1	4	2	2	1

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) ضریب تغییرات این داده جدید را بدست آورید (20)

X_i	111	113	112	114	117
f_i	1	4	2	2	1

در داده های فوق هر یک از داده جدید باندازه 100 واحد بیش از داده قدیم است

در خصوص X ها خود مسئله عنوان کرده که

$$\bar{x} = 13.2 \quad \sigma_x = \sqrt{2.36} = 1.536$$

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

$$\mu \pm \sigma = 13.2 \pm 1.536 = 11.664 \quad \text{تا} \quad 14.736$$

یعنی بیش از 68% داده ها بین 11.664 تا 14.736 میباشد

$$\mu \pm 2\sigma = 13.2 \pm (2 * 1.536) = 13.2 \pm 3.072 = 10.128 \text{ تا } 16.272$$

یعنی بیش از 96٪ داده ها بین 10.128 تا 16.272 میباشند
طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 100$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 100 = 13.2 + 100 \rightarrow \bar{y} = 113.2$$

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

$$\sigma_y = \sigma_x = \sqrt{2.36} = 1.536 \rightarrow \sigma_y = 1.536$$

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

$$\rho = \frac{\sigma}{\mu} = \frac{1.536}{113.2} = 0.0136 = 1.3\%$$

یعنی پراکندگی داده ها کم 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

6) نمرات دانشجویان بشرح زیر است داده ها را در جدول زیر مرتب کنید. الف) مد(نما) داده ها چه مقدار است (چرا) ب). میانه داده ها چه مقدار است؟ ج) دهک دوم را محاسبه کنید. د) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید د) ضریب تغییرات را بدست آورید. (20نمره)

نمرات: 14 و 12 و 12 و 14 و 15 و 14 و 16 و 16 و 11 و 14

نمره = x					
تعداد = f					

حل:

x داده	11	12	14	15	16
f فراوانی	1	2	4	1	2
F فراوانی تجمعی	1	3	7	8	10
r = $\frac{f_i}{\sum f_i}$ فراوانی نسبی	1/10=10%	2/10=20%	4/10=40%	1/10=10%	2/10=20%
R = $\frac{\sum f_{i_i}}{\sum f_i}$ فراوانی نسبی تجمعی	1/10=10%	3/10=30%	7/10=70%	8/10=90%	10/10=100%

مد داده ای بمقدار 14 میباشد زیرا بیشترین تعداد را دارد

$$M=14$$

میانه داده ها

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

دهک دوم

$$Q \sum f_i = \frac{2}{10}(1 + 2 + 4 + 1 + 2) = \frac{20}{10} = 2 \rightarrow \text{در سطر } F_i \rightarrow 2+ \rightarrow F_i=3 \rightarrow m=12$$

یعنی دو دهم داده ها 12 یا کمتر از 12 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(11 * 1) + (12 * 2) + (14 * 4) + (15 * 1) + (16 * 2)}{1 + 2 + 4 + 1 + 2} = \frac{138}{10} = 13.8$$

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

$$\sigma^2 = \frac{(11 - 13.8)^2 * 1 + (12 - 13.8)^2 * 2 + (14 - 13.8)^2 * 4 + (15 - 13.8)^2 * 1 + (16 - 13.8)^2 * 2}{1 + 2 + 4 + 1 + 2} = 2.56 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{2.56} = 1.6 \text{ انحراف معیار}$$

جدول در بالا تکمیل شد

نسبی r یعنی اینکه مثلاً از داده به مقدار 14 بتعداد 4 تا داریم یعنی 40٪ داده ها 14 هستند

نسبی تجمعی R یعنی اینکه مثلاً از داده به مقدار 14 و کمتر از 14 بتعداد 7 تا داریم یعنی 70٪ داده ها 14 یا کمتر از 14 هستند

ضریب تغییرات

$$\rho = \frac{\sigma}{\mu} = \frac{1.6}{13.8} = 0.1159 = 11.59\%$$

(7) تعدادی افراد بالای 70 سال میزان حافظه بلندمدت آنها را اندازه گرفتیم جداول زیر حاصل شد

X=میزان سن بالای 70 سال	3	4	4	5	4	5
Y=حافظه	11	14	12	12	11	14

الف) ضریب همبستگی پیرسون بین دو سری عدد X و Y را بدست آورید و در خصوص علامت این ضریب و مقدار آن توضیح بنویسید (ب) ضریب همبستگی اسپیرمن بین دو سری عدد X و Y را بدست آورید و در خصوص علامت این ضریب و مقدار آن توضیح بنویسید (20 نمره)

$$\rho = \frac{Cov(x,y)}{\sigma_x \cdot \sigma_y} = \frac{\bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}$$

$$r_s = 1 - \frac{6 \sum_{i=1}^n D^2}{n(n^2 - 1)}$$

							جمع= sum	sum/sigma_f				
X_i	۳	۴	۴	۴	۵	۵	25	4.17	6	میانگین x	4.17	p_x 0.16
Y_i	۱۱	۱۱	۱۲	۱۴	۱۲	۱۴	74	12.33	6	میانگین y	12.33	p_y 0.10
$X_i * Y_i$	33	44	48	56	60	70	311	51.83	6	میانگین xy	51.83	51.39
$(x_i - \bar{x})^2$	1.36	0.03	0.03	0.03	0.69	0.69	2.83	0.47	6	انحراف معیار x	0.69	
$(y_i - \bar{y})^2$	1.78	1.78	0.11	2.78	0.11	2.78	9.33	1.56	6	انحراف معیار y	1.25	
											Cov(x,y)	0.44
											$\rho(x,y)$	0.52

ضریب همبستگی پیرسون = $+0.52$ یعنی همبستگی مستقیم و متوسط

X_i	۳	۴	۴	۴	۵	۵
R_x	1	2	3	4	5	6
r_x	1	3	3	3	5.5	5.5

0.63

Y_i	۱۱	۱۱	۱۲	۱۴	۱۲	۱۴
R_y	1	2	3	5	4	6
r_y	1.5	1.5	3.5	5.5	3.5	5.5
$d=(r_x-r_y)$	-0.5	1.5	-0.5	-2.5	2	0
d^2	0.3	2.3	0.3	6.3	4	0

ضریب همبستگی اسپیرمن = $+0.63$ یعنی همبستگی مستقیم و متوسط

اسپیرمن	حاصل کرد	مخرج صورت
0.6	0.37	78 210