

سوالات و جواب امتحان میانترم آمار توصیفی 11-09-98 مدت 80 دقیقه -

سوالات 1 تا 4 مشترک و سوال 5 و 6 و 7 مربوط به یک گروه و سوالات 8 و 9 و 10 مربوط به گروه دیگر است
اگر در جایی غلط دیدید خودتان تصحیح کنید و بنویسید غلط

1) جمعیت چیست نمونه چیست (1نمره)

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

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.

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

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 لاینکرت 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) دو مورد از انواع نمودارها نام ببرید و هر یک نیم خط توضیح دهید و شکل نمودار را ترسیم کنید (2نمره)



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



5) جدول سوال زیر را قبلا حل شده است و میانگین 14.3 و واریانس 5.81 میباشد .

X_i	10	12	14	15	18
f_i	1	2	2	3	2

الف) انحراف معیار را حساب کنید

ج) میانگین با انحراف معیار جمع و تفریق کنید و یکبار 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 ج میانگین و انحراف معیار جدول زیر را بشکل سریع محاسبه کنید(3نمره)

X_i	11	13	15	16	19
f_i	1	2	2	3	2

$$\sigma^2 = 5.81 \rightarrow \sigma = \sqrt{\sigma^2} = \sqrt{5.81} = 2.41$$

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

یعنی بیش از 68% داده ها بین 11.89 تا 16.71 میباشند $\mu \pm \sigma = 14.3 \pm 2.41 = 11.89$ تا 16.71

یعنی بیش از 96% داده ها بین 9.48 تا 19.12 میباشند $\mu \pm 2\sigma = 14.3 \pm (2 * 2.41) = 14.3 \pm 4.82 = 9.48$ تا 19.12

داده جدید و مقایسه با داده قدیم

$X =$ داده قدیم	X_i	10	12	14	15	18
فراوانی	f_i	1	1	2	3	2
$y =$ داده جدید	X_i	11	13	15	16	19

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

در داده های فوق هر یک از داده جدید باندازه 1 واحد بیش از داده قدیم است

در خصوص x ها در صفحه قبل بدست آوردیم که

$$\bar{x} = 14.3 \quad \sigma_x = 2.41$$

طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 1$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 1 = 14.3 + 1 \rightarrow \bar{y} = 15.3 \quad \sigma_y = \sigma_x \rightarrow \sigma_y = 2.41$$

میانگین داده های جدید یک واحد بیشتر از میانگین داده قدیم است و واریانس داده جدید مساوی واریانس داده قدیم است

6) نمرات دانشجویان بشرح زیر است داده‌ها را در جدول زیر مرتب کنید. الف) میانه داده‌ها چه مقدار است؟
 (چرا) ب) مد(نما) داده‌ها چه مقدار است. - ج) صدک 25 داده‌های را معلوم کنید - د) میانگین و واریانس
 و انحراف معیار نمرات دانشجویان را بدست آورید. ه) با تکمیل جدول، فراوانی نسبی و فراوانی نسبی
 تجمعی را محاسبه کنید و با یک مثال نسبی‌ها را توضیح دهید. نمرات: 14 و 12 و 11 و 18 و 14 و 11 و 12
 و 12 و 10 و 12 (5نمره)

داده	X_i	10	11	12	14	18
فراوانی	f_i	1	2	4	2	1
فراوانی تجمعی	F_i	1	3	7	9	10
فراوانی نسبی = $\frac{f_i}{\sum f_i}$	r_i	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{1}{10}$
فراوانی نسبی تجمعی	R_i	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{7}{10}$	$\frac{9}{10}$	$\frac{10}{10}$

مد 12 زیرا بیشترین تعداد را دارد

میانه

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 4 + 2 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=7 \rightarrow m=12$$

صدک 25

$$Q \sum f_i = \frac{25}{100}(1 + 2 + 4 + 2 + 1) = \frac{250}{100} = 2.5 \rightarrow \text{در سطر } F_i \rightarrow 2.5+ \rightarrow F_i=3 \rightarrow 11$$

25٪ داده‌ها 11 یا کمتر از 11 هستند

میانگین

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(10 * 1) + (11 * 2) + (12 * 4) + (14 * 2) + (18 * 1)}{1 + 2 + 4 + 2 + 1} = \frac{126}{10} = 12.6$$

واریانس

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

$$\sigma^2 = \frac{(10 - 12.6)^2 * 1 + (11 - 12.6)^2 * 2 + (12 - 12.6)^2 * 4 + (14 - 12.6)^2 * 2 + (18 - 12.6)^2 * 1}{1 + 2 + 4 + 2 + 1} = 4.64$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.64} = 2.15$$

بنابراین میانگین 12.6 واریانس 4.64 و انحراف معیار 2.15 میباشد

توضیح یک ستون از جدول فوق - مثلاً در ستون داده 11 :

f_i = فراوانی = تعداد 11 به میزان 2 تا میباشد

F_i = فراوانی تجمعی = تعداد داده 11 و کمتر از 11 به میزان 3 تا میباشد

r_i = فراوانی نسبی = دو دهم داده‌ها (یا 20٪ داده‌ها) از کل داده‌ها عدد 11 است

R_i = فراوانی نسبی تجمعی = سه دهم داده‌ها (یا 30٪ داده‌ها) 11 یا کمتر از 11 هستند

X_i	11
f_i	2
F_i	3
r_i	$\frac{2}{10}$
R_i	$\frac{3}{10}$

7) در یک منطقه میزان مبتلایان به ازدیاد قندخون در چند سال گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال 1398 چقدر میباشد (3نمره)

سال=X	1395	1396	1397	1398
میزان قند	3	5	6	??

حل : سال 96 را مبنای صفر قرار میدهیم تا راحتتر عملیات ضرب و تقسیم صورت پذیرد

سال=X	-1	0	+1	2
مصرف=Y	3	5	6	??

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

$$b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} = \frac{(3 * -1) + (5 * 0) + (6 * 1) - \frac{(-1 + 0 + 1) * (3 + 5 + 6)}{3}}{(-1)^2 + 0^2 + (1)^2 - \frac{(-1 + 0 + 1)^2}{3}}$$

$$b = \frac{-3 + 0 + 6 - \frac{0 + 14}{3}}{+1 + 0 + 1 - \frac{0}{3}} \quad b = \frac{3}{2} = 1.5$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-1 + 0 + 1}{3} \quad \bar{x} = 0$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{3 + 5 + 6}{3} \quad \bar{y} = \frac{14}{3}$$

$$\bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x} \rightarrow a = \frac{14}{3} - \frac{3}{2} * 0 \rightarrow a = \frac{14}{3} \rightarrow y = a + bx \rightarrow y = \frac{14}{3} + \frac{3}{2}x$$

$$1398 - 1396 = 2 \rightarrow y = \frac{14}{3} + \frac{3}{2}x \rightarrow y = \frac{14}{3} + \left(\frac{3}{2} * 2\right) = \frac{23}{3} = 7.66$$

پس برای سال 98 پیش بینی به مقدار 7.66 میشود

8) جدول سوال زیر را قبلا حل شده است و میانگین 14.1 و واریانس 6.29 میباشد .

X_i	11	12	14	15	16	20
f_i	1	3	2	2	1	1

الف) انحراف معیار را حساب کنید

ج) میانگین با انحراف معیار جمع و تفریق کنید و یکبار 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 ج میانگین و انحراف معیار جدول زیر را بشکل سریع محاسبه کنید(3نمره)

X_i	10	11	13	14	15	19
f_i	1	3	2	2	1	1

$$\sigma^2 = 6.29 \rightarrow \sigma = \sqrt{\sigma^2} = \sqrt{6.29} = 2.51$$

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

یعنی بیش از 68% داده ها بین 11.59 تا 16.61 میباشند

یعنی بیش از 96% داده ها بین 9.08 تا 19.12 میباشند

داده جدید و مقایسه با داده قدیم

$x =$ داده قدیم	X_i	11	12	14	15	16	20
فراوانی	f_i	1	3	2	2	1	1
$y =$ داده جدید	X_i	10	11	13	14	15	19

قانونی داریم که اگر دادهها به یک مقدار یکسان اضافه شوند میانگین هم بهمان مقدار اضافه میشود ولی واریانس ثابت میماند

در داده های فوق هر یک از داده جدید باندازه 1 واحد بیش از داده قدیم است

در خصوص x ها در صفحه قبل بدست آوردیم که

$$\bar{x} = 14.1 \quad \sigma_x = 2.51$$

طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i - 1$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} - 1 = 14.1 - 1 \rightarrow \bar{y} = 13.1 \quad \sigma_y = \sigma_x \rightarrow \sigma_y = 2.51$$

میانگین داده های جدید یک واحد کمتر از میانگین داده قدیم است و واریانس داده جدید مساوی واریانس داده قدیم است

9) نمرات دانشجویان بشرح زیر است داده‌ها را در جدول زیر مرتب کنید. الف) میانه داده‌ها چه مقدار است؟
 (چرا) ب) مد(نما) داده‌ها چه مقدار است. - ج) صدک 25 داده‌های را معلوم کنید - د) میانگین و واریانس
 و انحراف معیار نمرات دانشجویان را بدست آورید. ه) با تکمیل جدول، فراوانی نسبی و فراوانی نسبی
 تجمعی را محاسبه کنید و با یک مثال نسبی‌ها را توضیح دهید. نمرات: 16 و 14 و 13 و 20 و 16 و 13 و 14
 14 و 12 و 14 (5نمره)

داده	X_i	12	13	14	16	20
فراوانی	f_i	1	2	4	2	1
فراوانی تجمعی	F_i	1	3	7	9	10
فراوانی نسبی = $\frac{f_i}{\sum f_i}$	r_i	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{1}{10}$
فراوانی نسبی تجمعی	R_i	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{7}{10}$	$\frac{9}{10}$	$\frac{10}{10}$

مد 14 زیرا بیشترین تعداد را دارد

میانه

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

صدک 25

$$Q \sum f_i = \frac{25}{100}(1 + 2 + 4 + 2 + 1) = \frac{250}{100} = 2.5 \rightarrow \text{Fi در سطر } 2.5+ \rightarrow \text{Fi}=3 \rightarrow 13$$

25٪ داده‌ها 13 یا کمتر از 13 هستند

میانگین

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(12 * 1) + (13 * 2) + (14 * 4) + (16 * 2) + (20 * 1)}{1 + 2 + 4 + 2 + 1} = \frac{146}{10} = 14.6$$

واریانس

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f_i}{\sum f_i} = \frac{(12 - 14.6)^2 * 1 + (13 - 14.6)^2 * 2 + (14 - 14.6)^2 * 4 + (16 - 14.6)^2 * 2 + (20 - 14.6)^2 * 1}{1 + 2 + 4 + 2 + 1} = 4.64$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.64} = 2.15$$

بنابراین میانگین 14.6 واریانس 4.64 و انحراف معیار 2.15 میباشد

توضیح یک ستون از جدول فوق - مثلاً در ستون داده 13 :

f_i = فراوانی = تعداد 13 به میزان 2 تا میباشد

F_i = فراوانی تجمعی = تعداد داده 13 و کمتر از 13 به میزان 3 تا میباشد

r_i = فراوانی نسبی = دو دهم داده‌ها (یا 20٪ داده‌ها) از کل داده‌ها عدد 13 است

R_i = فراوانی نسبی تجمعی = سه دهم داده‌ها (یا 30٪ داده‌ها) 13 یا کمتر از 13 هستند

\bar{X}_i	11
f_i	2
F_i	3
r_i	$\frac{2}{10}$
R_i	$\frac{3}{10}$

10) در یک منطقه میزان مبتلایان به ازدیاد قندخون در چند سال گذشته بشرح ذیل میباشد معادله خط رگرسیون را نوشته پیش بینی سال 1398 چقدر میباشد (3نمره)

سال=X	1395	1396	1397	1398
میزان قند	2	3	5	??

حل : سال 96 را مبنای صفر قرار میدهیم تا راحتتر عملیات ضرب و تقسیم صورت پذیرد

سال=X	-1	0	+1	2
مصرف=Y	2	3	5	??

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

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

$$b = \frac{-3 + 0 + 5 - \frac{0 + 14}{3}}{+1 + 0 + 1 - \frac{0}{3}} \quad b = \frac{3}{2} = 1.5$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-1 + 0 + 1}{3} \quad \bar{x} = 0$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{2 + 3 + 5}{3} \quad \bar{y} = \frac{10}{3}$$

$$\bar{y} = a + b\bar{x} \rightarrow a = \bar{y} - b\bar{x} \rightarrow a = \frac{10}{3} - \frac{3}{2} * 0 \rightarrow a = \frac{10}{3} \rightarrow y = a + bx \rightarrow y = \frac{10}{3} + \frac{3}{2}x$$

$$1398 - 1396 = 2 \rightarrow y = \frac{10}{3} + \frac{3}{2}x \rightarrow y = \frac{10}{3} + \left(\frac{3}{2} * 2\right) = \frac{19}{3} = 6.33$$

پس برای سال 98 پیش بینی به مقدار 6.33 میشود