

امتحان پایانترم آمار زیستی 8 سوال در زیر هر سوال چندین سوال مطرح شده است - (ارزش امتحان پایانترم برای درس آمار ۱۲ نمره از ۲۰ نمره و برای کارگاه آمار ۵ نمره از ۲۰ نمره میا شد) - حتما در یک محل خالی در وسط هر برگ پاسخ نام و فامیل نوشته و با تکمیل جدول زیر در ابتدا و همه تصاویر پاسخها در یک فایل PDF ارسال کنید - استفاده از جزوه ممنوع مشورت با دیگران ممنوع - فقط ماشین حساب آزاد - مدت ۹۰ دقیقه که بعد از ۷۵ دقیقه شروع به بارگذاری کنید تا با مشکل اینترنت و اتمام وقت مواجه نشوید

سوال ۵	انحراف معیار	فراموشی توان دوم در واریانس	$\mu \pm \sigma$	$\mu \pm 2\sigma$	میانگین داده جدید	انحراف معیار داده جدید	ضریب تغییرات داده جدید
پاسخ	2.002	0	11.297	9.295	113.3	2.002	0.017=1.7%
			15.302	17.305			

سوال ۶	مود	میان	صدک ۲۵	میانگین	واریانس	انحراف معیار
پاسخ	14	14	12	14	4.4	2.1

سوال ۷	حداکثر ۱۵	بیشتر از ۱۹	مساوی ۱۷
پاسخ	0.15	0.15	0.2

سوال ۸	r_s
پاسخ	0.86

۱) نام دو نرم افزار در آمار که در این ترم تحصیلی آموزش دیدید را ذکر کنید (۵نمره)

SPSS (Statistical package for social science) SAS (Statistical Analysis Software)

Minitab Eviews Statistica Lisrel Expert choice
 NCSS Microfit GAMS STATA SmartPLS Amos
 Statgraph PASS(NCSS) R Excel Mathlab

۲) جمعیت چیست نمونه چیست (۵نمره)

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.

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

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 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 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

۴) چهار مورد از انواع نمودارها نام ببرید و هر یک نیم خط توضیح دهید و شکل نمودار را ترسیم کنید (۵نمره)



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

۵) جدول سوال زیر قبلا حل شده است و میانگین 13.3 و واریانس 4.01 میباشد. مجدد حل نکنید

X_i	10	11	12	15	16
f_i	1	1	3	4	1

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 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	110	111	112	115	116
f_i	1	1	3	4	1

$$\bar{x} = 13.3 \quad \sigma_x = \sqrt{4.01} = 2.0025$$

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

$$\mu \pm \sigma = 13.3 \pm 2.0025 = 11.2975 \quad \text{تا} \quad 15.3025$$

یعنی بیش از ۶۸٪ داده ها بین 11.297 تا 15.302 میباشند

$$\mu \pm 2\sigma = 13.3 \pm (2 * 2.0025) = 13.3 \pm 4.005 = 9.295 \quad \text{تا} \quad 17.305$$

یعنی بیش از ۹۶٪ داده ها بین 9.295 تا 17.305 میباشند
 طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 100$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 100 = 13.3 + 100 \rightarrow \bar{y} = 113.3$$

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

$$\sigma_y = \sigma_x = \sqrt{4.01} = 2.0025 \rightarrow \sigma_y = 2.0025$$

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

$$\rho = \frac{\sigma}{\mu} = \frac{2.0025}{113.3} = 0.0177 = 1.7\%$$

الان $\rho = 0.0177$ 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

۶) نمرات دانشجویان بشرح زیر است. الف) مد(نما) داده ها چه مقدار است (چرا) ب). میانه داده ها چه مقدار است؟ ج) صدک ۲۵ را محاسبه کنید. د) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید (۲۰نمره)

نمره = x	10	12	14	15	17
تعداد = f	1	2	3	2	2
F	1	3	6	8	10

مد داده ای بمقدار 14 میباشد زیرا بیشترین تعداد را دارد

$$M=14$$

میانه داده ها

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 3 + 2 + 2) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=6 \rightarrow m=14$$

یعنی نصف داده ها 14 یا کمتر از 14 هستند

صدک ۱۴ داده ها

$$Q \sum f_i = \frac{25}{100}(1 + 2 + 3 + 2 + 2) = \frac{25}{100} = 2.5 \rightarrow \text{در سطر } F_i \rightarrow 2.5+ \rightarrow F_i=3 \rightarrow x=12$$

یعنی 25% داده ها 12 یا کمتر از 12 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(10 * 1) + (12 * 2) + (14 * 3) + (15 * 2) + (17 * 2)}{1 + 2 + 3 + 2 + 2} = \frac{140}{10} = 14$$

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

$$\sigma^2 = \frac{(10 - 14)^2 * 1 + (12 - 14)^2 * 2 + (14 - 14)^2 * 3 + (15 - 14)^2 * 2 + (17 - 14)^2 * 2}{1 + 2 + 3 + 2 + 2} = 4.4 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.4} = 2.1 \text{ انحراف معیار}$$

۷) یک دانشجو ۱۱۰ واحد درسی را با توزیع نرمال و میانگین نمرات 17 و انحراف معیار ۲ گذرانده است، در یک درس در ترم بعدی. الف) احتمال اینکه نمره این دانشجو حداکثر 15 شود چقدر است؟ ب) احتمال اینکه نمره این دانشجو بیشتر از 19 شود چقدر است ج) احتمال اینکه نمره دقیقا 17 شود؟ (جدول توزیع نرمال استاندارد بشرح زیر در نظر بگیرید ۲۰ نمره)

$$p(z \leq 2) = 0.98, \quad p(z \leq 1.5) = 0.93, \quad p(z \leq 1) = 0.85, \quad p(z \leq 0.5) = 0.7, \quad p(z \leq 0) = 0.5$$

$$p(z \leq -2) = 0.02, \quad p(z \leq -1.5) = 0.07, \quad p(z \leq -1) = 0.15, \quad p(z \leq -0.5) = 0.3$$

$$p(z \leq 2) = 0.98, \quad p(z \leq 1.5) = 0.93, \quad p(z \leq 1) = 0.85, \quad p(z \leq 0.5) = 0.7, \quad p(z \leq 0) = 0.5$$

$$p(z \leq -2) = 0.02, \quad p(z \leq -1.5) = 0.07, \quad p(z \leq -1) = 0.15, \quad p(z \leq -0.5) = 0.3$$

$$P(x \leq 15) = P\left(\frac{x - \mu}{\sigma} \leq \frac{15 - 17}{2}\right) = P(z \leq -1) = 0.15$$

$$P(x > 19) = 1 - P(x \leq 19) = 1 - P\left(\frac{x - \mu}{\sigma} \leq \frac{19 - 17}{2}\right) = 1 - P(z \leq 1) = 1 - 0.85 = 0.15$$

$$P(x = 17) = P(x \leq 17) - P(x \leq 16) = P\left(\frac{x - \mu}{\sigma} \leq \frac{17 - 17}{2}\right) - P\left(\frac{x - \mu}{\sigma} \leq \frac{16 - 17}{2}\right) = P(z \leq 0) - P(z \leq -0.5) = 0.5 - 0.3 = 0.2$$

۸) تعدادی افراد بالای ۷۰ سال میزان فشارخون آنها را اندازه گرفتیم جدول زیر حاصل شد

X= میزان سن بالای ۷۰ سال	3	4	4	5	5	5
Y= فشارخون	1	2	2	2	3	3

ضریب همبستگی اسپیرمن (رتبه‌ای) بین دو سری عدد X و Y را بدست آورید و در خصوص علامت این ضریب و مقدار آن توضیح بنویسید (۲۰ نمره)

ضریب همبستگی اسپیرمن

x	3	4	4	5	5	5
rx	1	2	3	4	5	6
Rx	1	2.5	2.5	5	5	5

داده سری اول

رتبه سری اول

رتبه سری اول نرم شده

y	1	2	2	2	3	3
rY	1	2	3	4	5	6
Ry	1	3	3	3	5.5	5.5

داده سری دوم

رتبه سری دوم

رتبه سری دوم نرم شده

Rx-Ry	0	-0.5	-0.5	2	-0.5	-0.5
(Rx-Ry) ²	0	0.25	0.25	4	0.25	0.25

$$r_s = 1 - \frac{6 \sum_{i=1}^n D^2}{n(n^2 - 1)} = 1 - \frac{6 * \{0^2 + (-0.5)^2 + (-0.5)^2 + (2)^2 + (-0.5)^2 + (-0.5)^2\}}{6(6^2 - 1)} = 1 - \frac{6 * 5}{210}$$

$$= 1 - 0.1429 = +0.8571$$

ضریب همبستگی اسپیرمن +0.857 شد یعنی دو سری داده وابستگی زیاد در جهت مستقیم دارند