کاربرد آمار در محیط زیست

آمار توصيفي

آمار چیست؟ آمار علم جمع آورى؛ تنظيم؛ توصيف؛ استباط و تحليل و تفسير داده است . موضوع آمار چیست؟ موضوع آمار عبارتست از هنر و علم جمع آوری، تعبیر و تجزیه و تحلیل داده ها و استخراج نتایج منطقی در مورد پدیده های تحت بررسی. اهداف علم آمارچیست؟ هدف اول: توصيف ساده، صريح و قابل فهم مشاهده ها كه معمولاً در نمونه مناسبي از جامعه صورت مي گيرد. هدف دوم: تعميم نتايج مشاهده هاي مزبور به جامعه اي كه نمونه از آن برگرفته شده است. ابزار های اصلی علم آمارکدامند ؟ ۱- آمار توصیفی که ویژگی های کلی تعدادی از داده ها را (که معمولاً به نمونه تعلق دارند) در قالب نمودار؛ جداول و یا یک عدد بیان می کند. آمار توصیفی به عنوان یک ابزار مفید از اهداف اولیه علم آمار است. ۲- آمار استنباطی که بر اساس ویژگیهای مشاهده شده در نمونه، ویژگی های جامعه را برآورد می کند. آمار استنباطی شامل ابزار های مفید برای هدف دوم علم آمار است. آمار توصيفى: آمار توصیفی شامل ابزار بسیار مفیدی است که به ساده گی میتوان داده های آماری را تفسیر؛ توصیف و تحلیل نماییم. در این بخش این ابزار ها به طور مختصر قابل استفاده معرفی و علاوه بر این روش اجرای این ابزار با استفاده از نرم افزار SPSS برای چندین نمونه داده تحلیل می شود. ۱-جداول آماري: الف: جدول فروانی ساده: این جدول برای داده های گسسته؛ کیفی و پیوسته با انواع کم رسم می شود. ب: جدول فراوانی طبقه بندی: برای داده های پیوسته، گسسته با انواع زیاد رسم می شود. رسم جداوی آماری در نرم افزار SPSS

۱- رسم جدول فراوانی ساده

Step 1. Start the SPSS software, choose type in data and enter your data or open the data file.

Step2. In the SPSS data editor window, choose Analyze Then Descriptive Statistics, then Frequencies to open the Frequencies dialog box.

Step3. In the Frequencies dialog box: select the variable name of your data and transfer it into the variables: check the display frequency table box and the Ok.

۲- رسم جدول فراوانی طبقه ای

Step 1. In the SPSS data editor window choose Transform, then Record, Into Different variables to open the Record into Different variables dialog box.

کاربرد آمار در محیط زیست دانشگاه فردوسی مشهد. ۹۹–۹۸

Step2. In the Record into Different variables dialog box select. variable name of your data and transfer it into the Input variable-output variable box.

Type class in the name: box and then click on change, then old and new values, to open the Record Different variables: Old and New values dialog box.

Step 3. In the Record into Different variables. Old and New values select Range and type c1=min x-i-a to the left of through and c1+h to the right of through. Check the Output variables are strings box and type 11 in the width. Also type ci-c(i+1) in the box for New value. Then click on Add and repeat this action until all the range are set.

Then Continue, then Analyze, then Frequencies, then the file of class transfer into Variables, check Display frequency tables and then OK.

۲-نمودارهای آماری الف: نمودار هیستوگرام: برای بررسی شکل توزیع به خصوص نرمال بودن یا غیر نرمال بودن داده ها رسم می شود. ب:نمودار جعبه ای: یک نمایش تصویری برای بررسی تقارن یا عدم تقاراز ماکزیمم؛ مینیمم و چارکهای توزیع بیان می کند معمولا برای مقایسه حداقل دو سری داده استفاده میشود.

۱-مرکز داده ها (میانه): خط میانی جعبه ۲-تغییرات یا پرکنش (دامنه میان چارکی) ارتفاع جعبه ۳- چولگی ( چولگی چارکی) اندازه نسبی نیم جعبه ها ۴-وجود یا عدم وجود مقادیر غیر عادی( مقادیر بیرون و یا خیلی بیرون) جارکی) اندازه نسبی نیم جعبه ها ۴-وجود یا عدم وجود مقادیر غیر عادی( مقادیر بیرون و یا خیلی بیرون) ج : نمودار احتمال: این نمودار رسم می شود تا مشخص نماید؛ چگونه داده ها بخوبی توسط توزیع جامعه مانند( نرمال- ج : نمودار احتمال: این نمودار رسم می شود تا مشخص نماید؛ چگونه داده ها بخوبی توسط توزیع جامعه مانند( نرمال- ج : نمودار این نمودار رسم می شود. در این نمودار انحراف از خط راست دلالت بر عدم برازش داده ها به توزیع جامعه مانند( برمال- لوگ نرمال و یا گاما) برازس داده می شود. در این نمودار انحراف از خط راست دلالت بر عدم برازش داده ها به توزیع جامعه مانند( جامعه می باشد.

د : نمودار Q-Q . برای مقایسه مستقیم دو سری داده رسم می شود. چنک های داده ها در یک نمودار رسم می شود درصورتی که نمدار پراکنش اطراف یک خط راست با شد یکسان بودن توزیع دو سری داده تایید می شود. کاربری مشابه نمودار احتمال نیز دارد که در آن یک محور چنک های داده ها و یک محور چند کهای توزیع جامعه است. ذ : نمودار ساقه و برگ: یک نمودار میله ای ستونی از داده ها را ارایه می دهد و داده های واقعی در هر لحظه در اختیار پژوهشگر هستند به عبارتی دیگر تصویر واقعی از توزیع داده ها را نشان می دهد. ر : نمودار دایره ای: برای تحلیل توصیفی داده های کیفی

۱- رسم نمودار های آماری برای جداول فراوانی ساده

Analyze---Descriptive Statistics---Frequencies----Transfer file of data into Variables, then click Charts and then select your charts based on frequencies or Percentages.

۲- رسم نمودار ساقه و برگ

Analyze-----Descriptive Statistics----Explore-----Transfer file of data into Dependent list, then select Plots for display. Click on Plots to open the Explore: Plots dialog box. Select None for Boxplots and Stem-and-leaf for Descriptive and then click on continue to return to the Explorer dialog box. Then click on OK.

۳-( شاخص های مرکزی) شاخصهای مرکزی یا تمرکز برای مطالعه مرکز داده ها که در تحلیل های آماری نقش مهمی دارد بررسی می شوند. در ادامه شاخصهای معروف و پر کاربرد بیان می شود. میانگین حسابی (Mean) :

به نقطه تعادل یا مرکز ثقل توزیع ، در داده هایی که بصورت منظم بر روی یک محور ردیف شده باشند ، میانگین

$$\overline{X} = \frac{\sum_{i=1}^{n} X_{i}}{n} = \frac{X_{1} + X_{2} + \dots + X_{n}}{n}$$
 (Mean ) اطلاق می شود

میانه (Median) :

مقداری است که نیمی (۵۰٪) مشاهدات کمتر از آن مقدار باشد. اگر n فرد باشد خواهیم داشت :

میانه 
$$Md = rac{X_{(rac{n}{2})} + X_{(rac{n}{2}+1)}}{2}$$

میانه =  $Md = X_{(\frac{n+1}{2})}$ 

## مُد (نما) (Mode) :

به مقداری گفته می شود که در میان سایر مقادیر توزیع ، بیشترین تکرار را داشته باشد ، مد را با Mo نشان می دهند. اشکالی که در این شاخص وجود دارد این است که امکان دارد در یک جامعه بیش از یک مُد داشته باشیم.

چار کها (Quartiles) : چار کها مقادیری هستند که داده ها را به چهار قسمت مساوی تقسیم میکنند. چارک اول: ۲۵ درصد داده ها کوچکتر یا مساوی آن هستند. با Q نشان میدهند چارک دوم: ۵۰ درصد داده ها کوچکتر یا مساوی آن هستند. با Q2 نشان می دهند. چارک سوم: ۲۵ درصد داده ها کوچکتر یا مساوی آن هستند. با Q3 نشان میدهند. \*-شاخصهای پراکندگی : کمیتهایی هستند که میزان تراکم یا عدم تراکم داده ها را اندازه گیری و بیان می کنند. دامنه تغییرات: (Range) دامنه تغییرات: (و مقادر کمترین مقدار مشاهده با بیشترین مقدار مشاهده. از معایب دامنه تغییرات این است که فقط از دو مقدار کمترین و بیشترین تأثیر میپذیرد و بقیه مشاهدات در نشان دادن وضعیت پراکندگی نقشی ندارند.  $R = X_{max} - X_{min}$ 

متوسط انحرافات از میانگین : این شاخـص از تقسیم مجموع قدر مطلـق انحـرافات تک تک مشاهدات از میانگین شان بر تعداد مشاهدات بدست می آید.

کاربرد آمار در محیط زیست

$$M.D = \frac{\sum \left| x_i - \overline{x} \right|}{n}$$

واریانس (Variance) واریانس داده ها که میزان تراکم داده حول مرکز ثقل را اندازه گیری و بیان می نماید؛ به دو صورت تعریف و در کاربردها استفاده می شود.

- I این رابطه یک تخمین نااریب برای واریانس جامعه ارایه می دهد.  

$$S_{1}^{\ 2} = \frac{n \sum x_{i}^{2} - (\sum x_{i})^{2}}{n(n-1)} = \frac{\sum (x_{i} - \bar{x})^{2}}{n-1}$$
- S<sub>2</sub><sup>2</sup> =  $\frac{\sum (x_{i} - \bar{x})^{2}}{n} = \frac{1}{n} \left\{ \sum_{i=1}^{n} x_{i}^{2} - n \bar{x}^{2} \right\}$ 

مقادير

انحراف معيار (Standard Deviation)

این معیار مشابه واریانس تعبیر می شود با این تفاوت که واحد انحراف معیار همان واحد داده هاست. به صورت زیر تعریف می شود

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} \qquad \qquad S = \sqrt{S^2}$$

اختلاف میان چارکی(Inter Quintile Range) این کمیت با حذف داده های کوچک ۲۵ درصد پایینی و حذف داده های بزرگ ۲۵ درصد بالایی به صورت  $IQR = Q_3 - Q_1$  محاسبه می شود. در واقع ۵۰ درصد داده ها در دامنه ( $Q_1, Q_3$ ) قرار میگیرند. مقادیر کوچک معیار اختلاف تراکم داده ها و مقادیر بزرگ پراگنده گی داده ها اطراف میانه را نشان می دهند. توجه: اگر فاصله های ( $Q_1, Q_2$ ) و ( $Q_2, Q_3$ ) تقریبا برابر باشند توزیع متقارن است. اگر فاصله ( $Q_1, Q_2$ ) کوچکتر از ( $Q_2, Q_3$ ) باشد توزیع نامتقارن و دارای چولگی مثبت و دم راست کشیده می باشد. اگر فاصله ( $Q_1, Q_2$ ) بزرگتر از ( $Q_2, Q_3$ ) باشد توزیع نامتقارن و دارای چولگی مثبت و دم راست کشیده می باشد. اگر فاصله ( $Q_1, Q_2$ ) بزرگتر از ( $Q_2, Q_3$ ) باشد توزیع نامتقارن و دارای چولگی مثبت و دم راست کشیده می باشد. اگر فاصله ( $Q_1, Q_2$ ) بزرگتر از ( $Q_2, Q_3$ ) باشد توزیع نامتقارن و دارای چولگی مثبت و دم راست کشیده می باشد. اگر فاصله ( $Q_1, Q_2$ ) بزرگتر از ( $Q_2, Q_3$ ) باشد توزیع نامتقارن و دارای چولگی مثبت و دم راست کشیده می باشد. اگر فاصله ( $Q_1, Q_2$ ) منفی و دم چپ کشیده می باشد.

$$1 - aralloi (i control i):$$
 $ac = aulib = aulib iic

  $7 - \varphi = 0$ 
 $ac < aulib < aulib < aulib < aulib 
 $7 - \varphi = 0$ 
 $ac < aulib < aulib < aulib 
 $7 - \varphi = 0$ 
 $ac < aulib < aulib 
 $7 - \varphi = 0$ 
 $ac < aulib < aulib 
 $7 - \varphi = 0$ 
 $ac < au 
 $7 - \varphi = 0$$$$$$$$$$$$$$$$$$$$$$$$$$$ 

مفهوم چولگی :(Skewness)

اگر دم توزیع جامعه به سمت راست باشد ، توزیع را چوله به راست و در صورت عکس ، آن را چوله به چپ می نامند.

کاربرد آمار در محیط زیست



چوله به راست



تفسیر مقادیر 
$$SS$$
:  
تفسیر مقادیر  $SS$  :  $SK$  :  $SK$  :  $SK$  :  $N$  :  $SK$  :  $N$  :  $SK$  :  $N$  :  $SK$  :  $N$  :  $SK$  :  $SK$  :  $N$  :  $N$  :  $SK$  :

این ضریب که به صورت زیر تعریف می شود ارتفاع توزیع داده ها را نسبت به توزیع نرمال استاندارد اندازه گیری و بیان می نماید.

$$K = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^4 / S^4 - 3$$



کاربرد آمار در محیط زیست



Select Analyze----Descriptive Statistics-----Frequencies-----Transfer file of data into Variables, select Statistics and then choose in Dialog box of Statistics then select Quartiels-Centeral Tendency—Dispersion---and Skew ness, Kurtosis.

۴-شاخصهای مرکزی و پراکنده گی برای داده های جدول فراوانی طبقه ای در نرم افزار Now, you have variables class midpoint and frequency in SPSS Data Editor;

Select Data in menu----weight cases---- check weighted cases by and then Transfer frequency into weighted cases by then click OK.

Choose Analyze-----Desc.Stat-----Frequencies-----In dialog box of frequencies transfer midpoint into Variables-----click Statistics: In this dialog box check Values are group midpoints. you can choose Quartiles—Dispersion—central Tendency –Skewness and Kurtosis then click continue and then Ok..

کاربرد آمار در محیط زیست

پ: نرمال سازی: اگر داده های مورد مطالعه از توزیع نرمال پیروی نکنند در این صورت با استفاده از یک تبدیل مناسب داده به داده های با توزیع نرمال تبدیل میشوند. یک تبدیل کلی برای نرمال سازی داده ها تبدیل باکس و کاکس است که

به صورت  $y_i = \frac{x_i^{\lambda} - 1}{\lambda}, i = 1, 2, \dots, n$  به صورت  $y_i = \frac{x_i^{\lambda} - 1}{\lambda}$ 

λ	-1	-1/2	0	1/2	1
y <sub>i</sub>	$\frac{1}{x_i}$	$\frac{1}{\sqrt{x_i}}$	$\ln(x_i)$	$\sqrt{x_i}$	X <sub>i</sub>

داده های زیست محیطی

مثال ۱-هدف مطالعه تحلیل توصیفی داده های مربوط به دو پدیده زمین شناسی(تاثیر میران بارش بر ارتفاع رواناب). داده های استفاده شده در این پروژه مربوط به بارش وارتفاع رواناب در ۵۲ ایستگاه در یک ناحیه می باشد.واحد بارش و ارتفاع هر دو میلی متر می باشد. بارش و ارتفاع رواناب دو متغیری هستند که طی یک سال در یک حوضه آبریز اندازه گیری شده است. دادههای مورد نظر در جدول ۱ نشان داده شدهاند.

زير حوزه	ضريب رواناب	ميزان	ارتفاع رواناب
	متوسط	بارش	زير حوزه ها
		(mm)	(mm )
A11	0.100	202.97	20.37
A12	0.105	201.50	21.12
A131	0.115	221.48	25.56
A1321	0.128	239.90	30.79
A1322	0.125	230.58	28.78
A13231	0.120	232.39	27.92
A1323- INT	0.121	224.50	27.15
a1323	0.119	230.84	27.39
A132-INT	0.112	215.73	24.25
a132	0.102	209.21	21.35
A13-INT	0.109	211.56	23.15
a13	0.112	226.27	25.29
A1-INT	0.095	190.85	18.11

کاربرد آمار در محیط زیست دانشگاه فردوسی مشهد. ۹۹–۹۸

a1	0.105	216.79	22.68
A-INT	0.087	176.46	15.31
а	0.103	214.02	22.04
B111	0.123	213.27	26.22
B112	0.131	235.47	30.80
B113	0.129	230.81	29.73
B11-INT	0.106	201.63	21.35
b11	0.118	224.43	26.48
B1-INT	0.097	191.56	18.65
b1	0.114	219.18	24.90
B-INT	0.091	180.49	16.45
b	0.107	210.28	22.56
C1111	0.120	217.72	26.09
C111- INT	0.084	179.69	15.02
c111	0.097	192.32	18.74
C1121	0.106	201.60	21.45
C1122	0.145	248.52	36.12
C11231	0.110	216.71	23.81
C1123- INT	0.110	213.67	23.50
c1123	0.108	214.39	23.15
C112- INT	0.075	155.86	11.69
c112	0.103	208.05	21.41
C113	0.090	187.04	16.79
C11-INT	0.085	177.67	15.03
c11	0.092	196.27	18.04
C1-INT	0.085	176.22	15.07
c1	0.090	194.15	17.55
C-INT	0.083	168.15	13.95
С	0.090	192.66	17.26
AC-INT	0.079	164.16	13.03

کاربرد آمار در محیط زیست

ac	0.093	202.84	18.90
D1	0.102	192.63	19.62
D-INT	0.087	177.57	15.53
d	0.096	186.36	17.83
ACD-INT	0.093	176.74	16.35
acd	0.091	200.60	18.33
E	0.077	160.24	12.32
F	0.118	218.68	25.72
GALEH- INT	0.085	167.48	14.17
کل حوزہ	0.091	200.78	18.35

مثال ۲- نمونه های مورد مطالعه از پساب خروجی یک نیروگاه برق برداشت شده است. بررسی میزان کلسیم موجود در این پساب مورد بررسی قرار گرفته است.

نمونه	میلی	كلسيم(
	لر)	گرم بر ليت
١		۴۰,۴
٢		۳۹٫۵
٣		30,88
۴		۳۵,۸۲
۵		30,14
۶		۴۰,۴
۷		۳۹,۸
٨		۳۹٫۵
٩		۵۲,۴
۱.		۴۰,۳

کاربرد آمار در محیط زیست

مشهد. ۹۹–۹۸	دانشگاه فردوسی
-------------	----------------

377,74
36,1
۳۹,۶
۳۹,۶
41,٣
۳۶,۳

مثال ۳-داده های حاصل از آنالیز چند نمونه خاک یک دشت میزان عنصر روی را به شرح زیر اعلام می کند.مطلوب است: الف) شاخص های مرکزی و پراکندگی برای داده های زیر را بدست آورده و تحلیل کنید.

شماره نمونه	روی بر حسب	
	ميكروگرم بر	
	گرم	
١	۷۱/۰	
٢	17.10	
٣	590	
۴	۶۵۱	
۵	184	
9	477	
γ	۱۲۱/۵	
٨	१४/•	
٩	۱۲۸	
١.	۱۲۳	
11	407	
١٢	۵۵۳	
١٣	۳۸۲	
14	۱۲۰	
۱۵	۶٣/۰	
١۶	۶۶/۵	
١٧	۱۷۳/۵	

کاربرد آمار در محیط زیست دانشگاه فردوسی مشهد. ۹۹–۹۸

مثال ۴-داده های استفاده شده در این پروژه مربوط به بارش و دبی در ۳۶ ایستگاه در شهرستان قوچان می باشد.واحد بارش میلیمتر و واحد دبی متر مکعب بر ثانیه هست.بارش و دبی هر دو متغیرهایی هستند که مربوط به سال های ۱۳۴۵تا ۱۳۸۱ هستند.

الف- تحليل توصيفي هر يک از متغيرها را انجام دهيد. ب) نرمال بودن توزيع متغيرها را بررسي كنيد



	۱۸۸,۰۰	۷,۲۰
	141,00	19,08
	۱۸۰,۶۰	14,98
	۱۰۷,۰۰	11,91
	۲۵۰,۲۰	11,91
	189,70	۲۸,۱۵
	107,70	۱۳,۰۷
	777,7.	۳۵,۲۰
	100,70	۱۷٫۸۹
	۲۳۶٫۸۰	۳۲,۲۲
	١٧٠,٧٠	۱۸,۱۸
	٨۶,۵۰	۲۰,۲۲
	۱۲۰,۰۰	14,40
	۷۸,۷۰	۶,۲۰
	189,80	۱۱٫۵۸
	۶۲,۷۰	18,18
9	۲۸۲,۶۰	۳۲,۱۱
	۵۸٫۱۰	74,41
	۱۰۵,۶۰	۱۰,۰۸
	۸۵,۱۰	۲۵,۹۵

کاربرد آمار در محیط زیست

	· · · · · · · · · · · · · · · · · · ·	
	۷۰,۱۰	70,94
	۲۴۷,۸۰	۳۰,۰۴
	118,80	44,49
	119,40	٩,۴٨
	۲۱۶,۱۰	55,08
	19.,9.	۲۹,۹۳
	179,9.	۱۰٫۵۹
	171,80	50,85
	۵۷,۷۰	۱۹٫۸۵
	184,80	۵,۹۶
	۹۵,۱۰	۴,۰۱
	۱۴۰٫۵۰	28,20
	١٢٧,٢٠	10,40
	۱۵۹,۵۰	74,17
	۳۳۰٫۵۰	74,17
	۶۷,۱۰	18,87
شاخصهای مرکزی با بدست آمرید و نمودار هیستو گراه با بسم کنید.	المستعمر المستعمم مستعمر المستعمم مستعمر مستعمر مستعمم مستعمم مستعمم مستعمر مستعمر مستعمم مستمم مستعمم مستمم مستعمم مستعمم مستمممم مستممممم مممممممم مستممممممم مستمممممممم	مای دلی ۵ الثم
		0 ). 0

دسته	فراوانی
•- <b>۵</b>	٣
۵-۱۰	٨
۱۰-۱۵	١٣
10-70	۴
۲۰-۲۵	٣



Histogram

Mean =19.06 Std. Dev. =9.423 N =36



Debi Stem-and-Leaf Plot

Frequency	Stem a	& Leaf
1,00 4,00 9,00 6,00 5,00 6,00 3,00 1,00 1,00	0 0 1 1 2 2 3 3 3 4	<pre>4 . 5679 . 001113344 . 567899 . 00444 . 556689 . 022 . 5 . 4</pre>
Stem width: Each leaf:	10,	,00 case(s)

دانشگاه فردوسی مشهد. ۹۹–۹۸

کاربرد آمار در محیط زیست

Statistics







کاربرد آمار در محیط زیست

دانشگاه فردوسی مشهد. ۹۹–۹۸