

فصل اول

اندازه گیری و مدل سازی: بیان مسئله به زبان ریاضی را مدل سازی می گوئیم.

خطای اندازه گیری: تفاضل مقدار واقعی و مقدار اندازه گیری شده می باشد.

نکته: در مدل های اندازه گیری چون E_1 ها مقادیری کوچک هستند پس توان دوم آن ها یا جمع درجه ۲ به بالا را حذف می کنیم یعنی تقریباً $E_1 E_2$ صفر است.

فصل دوم

جامعه آماری: مجموعه ای از افراد یا اشیاء است که در باره ی اعضای آن می خواهیم موضوع یا موضوعاتی را مطالعه کنیم.

نمونه: زیرمجموعه ای از جامعه آماری است.

سرشماری:

اگر تمام افراد جامعه را مورد مطالعه قرار دهیم سرشماری کرده ایم.

مشکلات سرشماری عبارتند از:

(۱) در دسترس نبودن تمام اعضای جامعه

(۲) وقت گیر بودن دسترسی به تمام اعضای جامعه

(۳) گران تمام شدن بررسی تمام اعضای جامعه

(۴) از بین رفتن جامعه در برقی از مطالعات

نمونه تصادفی ساده:

اگر امکان انتخاب هر فرد به عنوان عضوی از نمونه امکان پذیر باشد و قبل از انتخاب نمونه، نتوانیم با اطمینان

بیش تر درباره ی عضو یا عدم حضور عده ای در نمونه قضاوت کنیم، نمونه تصادفی می باشد.

متغیرهای تصادفی: هر موضوعی که در عضوهای مختلف یک جامعه متفاوت باشد متغیر تصادفی می گویند.

نوع متغیرهای تصادفی

متغیر کمی: متغیرهایی هستند که قابل اندازه گیری نمی باشند. مانند گروه خونی افراد

پیوسته: به متغیری که اگر دو مقدار a و b بتواند اختیار کند هر مقدار بین آنها

را نیز بتواند اختیار کند کمی پیوسته گویند. مانند وزن افراد

نوع متغیرهای کمی

گسسته: به متغیری که پیوسته نباشد گسسته گویند.

مثال: تعداد روزهای بارانی در یک شهر

کیفی ترتیبی: اگر در متغیرهای کیفی نوعی ترتیب طبیعی وجود داشته

باشد کیفی ترتیبی گویند. مانند مراحل زندگی انسان (نوزادی - قبل از

کودکی - کودکی و ...)

نوع متغیرهای کیفی

کیفی اسمی: اگر متغیر کیفی ویژگی فوق را نداشته باشد

به آن کیفی اسمی گویند. مانند: گروه خونی افراد

فصل چهارم:

فراوانی مطلق: تعداد دفعات تکرار هر داده مانند X_i را فراوانی یا فراوانی مطلق گویند و با f_i نمایش می دهند.

فراوانی نسبی: اگر فراوانی مطلق هر دسته را بر مجموع فراوانی های مطلق تقسیم کنیم فراوانی نسبی به دست

می آید.

(۱) در یک جدول فراوانی مجموع فراوانی های نسبی همواره برابر یک است.

(۲) فراوانی نسبی هر طبقه بین ۰ و ۱ می باشد.

درصد فراوانی نسبی: اگر فراوانی نسبی هر طبقه را در ۱۰۰ ضرب کنیم درصد فراوانی نسبی به دست می آید.

درصد فراوانی نسبی $F_i \times 100 = \frac{f_i}{n} \times 100$

فراوانی تجمعی: اگر فراوانی مطلق هر دسته را با فراوانی های دسته های قبل جمع کنیم فراوانی تجمعی به دست می آید.

نکته: در هر جدول توزیع فراوانی، فراوانی تجمعی دسته ی آخر برابر تعداد کل فراوانی ها می باشد.

نکته: در هر جدول فراوانی، فراوانی تجمعی دسته ی اول برابر فراوانی مطلق همان دسته است.

فراوانی تجمعی نسبی: تعداد فراوانی تجمعی هر طبقه تقسیم بر تعداد کل داده ها را فراوانی تجمعی نسبی گویند.

درصد فراوانی تجمعی نسبی: اگر فراوانی تجمعی نسبی هر طبقه را در ۱۰۰ ضرب کنیم به دست می آید.

دامنه ی تغییرات: اختلاف بین بزرگ ترین و کوچک ترین داده ی آماری را دامنه ی تغییرات گویند و طبق فرمول

زیر $R = b - a$ محاسبه می شود.

کوچکترین داده $R = b - a$ دامنه ی تغییرات \leftarrow بزرگترین عدد

نکته: اگر همه اعداد برابر باشند دامنه تغییرات صفر است.

نکته: در یک جدول فراوانی تعداد دسته ها را با k و طول دسته را با C نمایش می دهند.

نکته: رابطه ی بین دامنه تغییرات، طول و تعداد دسته ها عبارت است از:

$$C = \frac{R}{k}$$

طول دسته: هر دسته از دو عدد تشکیل شده که عدد کوچک تر را k و عدد بزرگ تر را C نمایش می دهند.

میان آنها را طول دسته نامیده و با C نمایش می دهند.

مرکز دسته: نصف مجموع عد پائین و عد بالای هر دسته را مرکز دسته یا نشان دسته گویند.

حد بالا + حد پایین $X_i = \frac{\text{حد بالا} + \text{حد پایین}}{2}$

نکته: اختلاف دو مرکز دسته متوالی، برابر طول دسته ها می باشد.

فصل پنجم:

نمودارها:

(۱) نمودار میله ای:

نکته: از این نمودار پیش تر برای متغیرهای کمی گسسته و کیفی استفاده می کنند.

نکته: در این نمودار ترتیب قرار گرفتن میله ها اهمیت ندارد.

نکته: در این نمودار روی محور افقی نشان دسته (X_i) و روی محور عمودی فراوانی مطلق می باشد.

نمودار مستطیلی:

نکته: با این نمودار برای متغیرهای کمی پیوسته مناسب است.

نکته: روی محور افقی حدود دسته ها و روی محور عمودی فراوانی مطلق یا نسبی قرار می گیرد.

نمودار چند بر فراوانی:

اگر بخواهیم تغییرات متغیر را بهتر نشان دهیم از این نمودار استفاده می کنیم.

نکته: روی محور افقی مرکز دسته و روی محور عمودی فراوانی مطلق می باشد.

نکته: بهتر است دو دسته با فراوانی صفر به ابتدا و انتهای دسته ها در نمودار چند بر فراوانی اضافه کنیم.

نمودار تجمعی:

این نمودار هم به صورت میله ای و هم به صورت مستطیلی قابل رسم می باشد.

نکته: روی محور طول ها حدود دسته یا مرکز دسته و روی محور عرض ها فراوانی تجمعی می باشد.

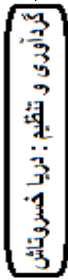
نمودار دایره ای:

این نمودار مناسب برای متغیرهای کیفی است.

نکته: در این روش مساحت دایره به قطعاتی تقسیم می شود که مساحت هر قطعه عبارت است از:

$$S_i = \frac{f_i}{n} \times 360$$

زاویه مربوطه $S_i = \frac{f_i}{n} \times 360$



نکته: اگر در نمودار دایره ای فراوانی داده‌ها را دو برابر کنیم زاویه مرکزی عوض نمی‌شود.

نکته: در نمودار دایره ای ترتیب قرار گرفتن نواحی مهم نمی‌باشد.

نمودار ساقه‌ی برگ:

از این نمودار زمانی که دامنه تغییرات کم است استفاده می‌کنیم.

نکته: این نمودار دو قسمت دارد ساقه و برگ

نکته: در این نمودار عدد را به دو قسمت تقسیم می‌کنیم، ساقه شامل یک یا چند رقم اولیه و برگ که بعد از آن نوشته می‌شود و همیشه شکل صعودی دارد، شامل ارقام باقیمانده است.

نکته: در قسمت برگ اگر اعداد تکراری باشند باید نوشته شوند.

نکته: در قسمت برگ اعداد باید صعودی نوشته شوند.

شاخص‌های مرکزی: به شاخص‌هایی که محل تمرکز داده‌ها را معرفی می‌کنند، شاخص‌های مرکزی گویند که عبارت است از میانگین، میانه و مد.

میانگین طبق این فرمول محاسبه می‌شود.

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{n}$$

میانه: میانه عددی است که از نصف داده‌ها کمتر و از نصف داده‌ها بیشتر است.

روش به دست آوردن میانه: ابتدا داده‌ها را مرتب می‌کنیم سپس داده‌ها را می‌شماریم.

الف) اگر تعداد داده‌ها فرد باشد، داده‌ای که در وسط قرار دارد میانه است.

ب) اگر تعداد داده‌ها زوج باشد، نصف مجموع دو داده‌ای که در وسط قرار گرفته‌اند میانه است.

چند نکته مهم درباره میانه:

۱) میانه منصفه به فرد است.

۲) میانه داده‌ها ممکن است از داده‌ها نباشد.

۳) اگر عدد ثابتی به تمام داده‌ها اضافه (کم) شود این عدد نیز به میانه اضافه می‌شود.

۴) اگر تمام داده‌ها را در عددی ضرب کنیم، میانه نیز در همان عدد ضرب می‌شود.

چارک‌ها:

اگر جامعه آماری را به چهار قسمت تقسیم کنیم چارک‌ها را به صورت زیر تعریف می‌کنیم.

۱) چارک اول: عددی که از $\frac{1}{4}$ داده‌ها بیشتر و از $\frac{3}{4}$ داده‌ها کمتر باشد. با Q_1 نمایش می‌دهند.

۲) چارک دوم: (میانه) عددی که از نصف داده‌ها بیشتر و از نصف داده‌ها کمتر باشد. با Q_2 نمایش می‌دهند.

۳) چارک سوم: عددی که از $\frac{3}{4}$ داده‌ها کمتر و از $\frac{1}{4}$ داده‌ها بیشتر باشد. با Q_3 نمایش می‌دهند.

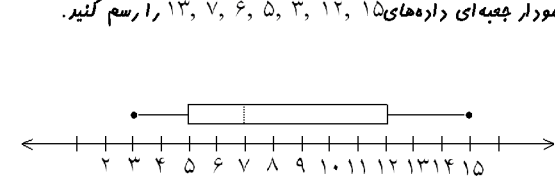
چارک اول: میانه‌ی نیمه اول داده‌ها را چارک اول گویند.

چارک سوم: میانه‌ی نیمه دوم داده‌ها را چارک سوم گویند.

نمودار جعبه‌ای: این نمودار پراکنندگی داده‌ها را بهتر نشان می‌دهد.

برای رسم این نمودار باید مینیمم و چارک اول و دوم و سوم و ماکسیمم داده‌ها را مشخص کنیم.

مثال: نمودار جعبه‌ای داده‌های ۱۵، ۱۲، ۳، ۵، ۶، ۷، ۱۳، ۱۴، ۱۵ را رسم کنید.



۳, ۵, ۶, ۷, ۱۲, ۱۳, ۱۵

min = ۳

max = ۱۵

$Q_1 = ۵$

$Q_3 = ۷$

$Q_3 = ۱۲$

مد: به آن داده‌ای گویند که بیشترین فراوانی را دارد.

مثال: در داده‌های مقابل مد کدام است. ۵، ۱، ۷، ۳، ۴، ۲

مد عدد ۴ می‌باشد.

نکته: مد منصفه به فرد نمی‌باشد یعنی در یک جامعه می‌توان چند مد داشت.

نکته: اگر به تمام داده‌ها عددی اضافه (کم) شود به مد نیز همان عدد اضافه می‌شود.

نکته: اگر در تمام داده‌ها عددی ضرب شود مد نیز در همان عدد ضرب می‌شود.

نکته: در جدول فراوانی سطر مد را آن سطری است که بیشترین فراوانی را دارد.

فصل هفتم

شاخص‌های پراکنندگی: میزان پراکنندگی در داده‌های آماری را نشان می‌دهد که عبارتند از: دامنه تغییرات،

واریانس، انحراف معیار، ضریب تغییرات

دامنه تغییرات: اختلاف بین کوچک‌ترین داده و بیشترین داده را دامنه تغییرات گویند.

$R = \max - \min$

انحراف از میانگین: مجموع قدر مطلق انحرافات داده‌ها از میانگین تقسیم بر تعداد داده‌ها را انحراف از میانگین

$$A = \frac{\sum |x_i - \bar{x}|}{n}$$

گویند که فرمول آن به صورت زیر است.

واریانس: واریانس برابر میانگین مجزور انحرافات از میانگین است و آن را با نماد δ^2 نمایش می‌دهند. بنابراین:

$$\delta^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

نکته: اگر مجموع مجزورات داده‌ها $\sum x_i^2$ و میانگین آنها \bar{x} داده شده باشد واریانس را می‌توان از رابطه

$$\delta^2 = \frac{\sum x_i^2}{n} - (\bar{x})^2$$

به دست آورد.

چند نکته درباره واریانس و انحراف معیار:

۱) اگر داده‌های آماری با عددی جمع یا تفریق شوند در واریانس و انحراف معیار تغییراتی ایجاد نمی‌شود. یعنی

$$\delta_{a+x}^2 = \delta_x^2$$

۲) اگر داده‌های آماری در عددی ضرب یا تقسیم شود، واریانس در توان دوم آن عدد ضرب می‌شود و انحراف

$$\delta_{ax}^2 = a^2 \delta_x^2 \quad \text{و} \quad \delta_{ax} = |a| \delta_x$$

معیار در قدر مطلق آن عدد ضرب می‌شود.

۳) اگر همه داده‌ها برابر باشند واریانس و انحراف معیار صفر هستند.

ضریب تغییرات: این شاخص پراکنندگی بدون واحد است و با نماد CV نمایش می‌دهند که عبارتست از تقسیم

$$CV = \frac{\delta}{\bar{x}}$$

انحراف معیار بر میانگین.

چند نکته درباره ضریب تغییرات:

۱) اگر همه داده‌ها با هم برابر باشند ضریب تغییرات صفر است.

۲) اگر همه داده‌ها در عددی ثابت و مثبت ضرب شوند ضریب تغییرات تغییر نمی‌کند.

۳) اگر همه داده‌ها با عددی جمع کنیم میانگین جدید بزرگ‌تر شده، در نتیجه ضریب تغییرات کم می‌شود.

۴) ضریب تغییرات میزان پراکنندگی داده‌ها را به ازای یک واحد میانگین نشان می‌دهد.

Website: <http://khosrotash.ir/>

Amoodarya@yahoo.com

daryakhosrotash@mehar.sharif.ir

خلاصه مطالب درس آمار

رودبازار کیلیبک می‌گوید: «من شش دوست صادق دارم که در خدمت من هستند.

آنها تمام چیزهایی را که من هم اکنون می‌دانم، به من آموخته‌اند. اسم این دوستان عبارتست از:

«کجا، چه، چه وقت، چرا، چگونه، چه کسی»

KHOSROTASH

موسسه علمی تلاشی
ابتدای شهرک مهر
33486003-4
09373324234