بسمه تعالی

موضوع : مبانی آمار و احتمال برای ورود به علم داده و تحلیل داده

بخش اول: آمار چیست و چرا مهم است؟

- آمار چیست؟
- تفاوت آمار و احتمالات

در این پارت از اموزش به بررسی **آما و احتمالات / کاربرد های آنها و تفاوت میان این دو** می پردازیم .

تقریباً هر روز در اخبار، شبکه‌های اجتماعی یا حتی محیط کار با جمله‌هایی مانند این مواجه می‌شویم:

- احتمال بارندگی فردا ۸۰ درصد است .
- میانگین درآمد خانوارها افزایش پیدا کرده است .
- احتمال موفقیت یک داروی جدید ۹۵ درصد است .
- احتمال وقوع جنگ
  احتمال وقوع زمین لرزه

اما آیا تا به حال از خودتان پرسیده‌اید که این اعداد چگونه به دست آمده‌اند؟

آیا همه این اعداد از علم آمار استخراج شده‌اند؟

یا بعضی از آن‌ها مربوط به علم احتمال هستند؟

در این جلسه قرار است دقیقاً به همین سؤال پاسخ دهیم.

آمار چیست؟

قبل از اینکه وارد مباحث تخصصی شویم، ابتدا باید بدانیم اصلاً آمار چیست.

آمار علمی است که به ما کمک می‌کند داده‌ها را جمع‌آوری کنیم، آن‌ها را سازماندهی کنیم، خلاصه کنیم، تحلیل کنیم و در نهایت از آن‌ها نتیجه‌گیری کنیم.

به عبارت ساده‌تر، هر جا داده وجود داشته باشد، آمار نیز حضور دارد.

فرض کنید مدیر یک کارخانه هستید و می‌خواهید بدانید:

- میانگین تولید روزانه چقدر است؟

- چند درصد محصولات معیوب هستند؟
- کدام خط تولید عملکرد بهتری دارد؟

تمام این پرسش‌ها با استفاده از علم آمار پاسخ داده می‌شوند.

کاربردهای آمار:

شاید تصور کنید آمار فقط مخصوص رشته آمار است؛ اما واقعیت این است که تقریباً هیچ علمی بدون آمار قابل تصور نیست.

برای مثال:

در **پزشکی** از آمار برای بررسی اثربخشی داروها استفاده می‌شود.

در **مدیریت** برای تحلیل عملکرد سازمان و تصمیم‌گیری.

در **مهندسی صنایع** برای کنترل کیفیت، برنامه‌ریزی تولید و بهبود فرآیندها.

در **اقتصاد** برای تحلیل نرخ تورم، رشد اقتصادی و بیکاری.

در **روان‌شناسی و علوم تربیتی** برای تحلیل پرسشنامه‌ها و آزمون‌های پژوهشی.

حتی در زندگی روزمره نیز هنگام تصمیم‌گیری، ناخودآگاه از مفاهیم آماری استفاده می‌کنیم.

احتمال چیست؟

حال سؤال مهم این است که احتمال چیست؟

اگر آمار به داده‌های گذشته نگاه می‌کند،

احتمال درباره اتفاق‌هایی صحبت می‌کند که هنوز رخ نداده‌اند.

احتمال میزان امکان وقوع یک رویداد را اندازه‌گیری می‌کند.

برای مثال:

اگر یک سکه سالم را پرتاب کنیم،

دو حالت ممکن وجود دارد:

- شیر
- خط

از آنجا که هر دو حالت شانس یکسانی دارند،

احتمال وقوع هر کدام برابر با ۰٫۵ یا ۵۰ درصد است.

یا اگر یک تاس سالم را پرتاب کنیم،

احتمال آمدن عدد ۶ برابر است با یک ششم.

بنابراین احتمال قبل از وقوع رویداد درباره میزان شانس وقوع آن صحبت می‌کند.

تفاوت آمار و احتمال:

اکنون به مهم‌ترین بخش این جلسه می‌رسیم.

خیلی از افراد تصور می‌کنند آمار و احتمال یک مفهوم هستند، در حالی که این دو مکمل یکدیگرند اما یکسان نیستند.

آمار از داده‌های موجود شروع می‌کند و به نتیجه می‌رسد.

یعنی:

داده ← تحلیل ← نتیجه

برای مثال:

نمرات ۵۰ دانشجو را جمع‌آوری می‌کنیم،

میانگین آن‌ها را محاسبه می‌کنیم،

و نتیجه می‌گیریم که عملکرد کلاس چگونه بوده است.

اما **احتمال** دقیقاً برعکس عمل می‌کند.

ابتدا یک مدل یا فرض درباره یک پدیده داریم،

سپس احتمال وقوع رویدادهای مختلف را محاسبه می‌کنیم.

یعنی:

مدل ← احتمال ← پیش‌بینی

یک مثال ساده

فرض کنید می‌خواهیم وضعیت بارندگی را بررسی کنیم.

اگر اطلاعات ۳۰ سال گذشته را جمع‌آوری و تحلیل کنیم،

در حال انجام یک مطالعه آماری هستیم.

اما اگر بگوییم:

(احتمال بارش فردا ۷۰ درصد است)

در واقع وارد حوزه احتمال شده‌ایم.

پس:

آمار گذشته را تحلیل می‌کند.

احتمال آینده را پیش‌بینی می‌کند.

ارتباط آمار و احتمال

در عمل، این دو علم کاملاً به هم وابسته‌اند.

ابتدا با استفاده از احتمال، مدل‌های ریاضی ساخته می‌شوند.

سپس با استفاده از آمار بررسی می‌کنیم که آیا داده‌های واقعی با آن مدل‌ها سازگار هستند یا خیر.

به همین دلیل در بسیاری از کتاب‌های دانشگاهی، این دو علم در کنار هم تدریس می‌شوند.