

به نام خداوند بخشنده مهربان

روش‌های آماری و اقتصادسنجی

در تحلیل و مدل‌سازی داده‌های حمل و نقلی

محمد مهدی بشارتی

besharati@iut.ac.ir



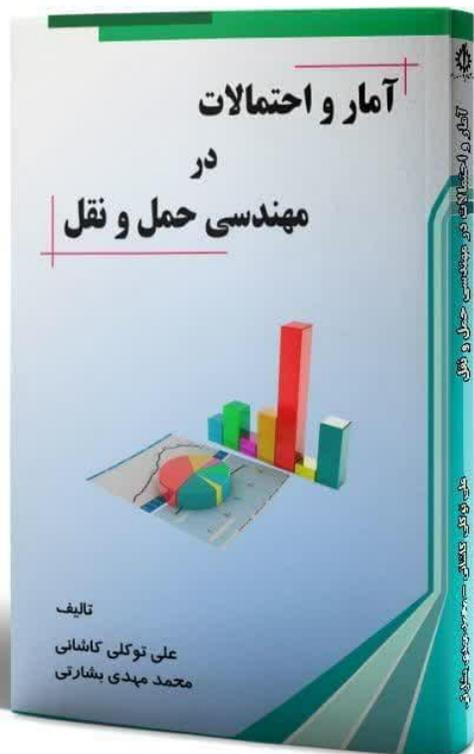
منابع فارسی

- ۱- «مبانی اقتصادسنجی»، دامودار گجراتی، دکتر حمید ابریشمی (مترجم)، دانشگاه تهران
- ۲- «مبانی احتمالات و آمار مهندسی». مجید ایوزیان، ابوالفضل واقفی، حمید اسماعیلی؛ انتشارات ترمه
- ۳- «آمار ریاضی و کاربردهای آن». جان فروند؛ محمد قاسم وحیدی اصل (مترجم)، علی عمیدی (مترجم)؛ نشر دانشگاهی
- ۴- «مقدمه ای بر احتمالات و آمار کاربردی». رونالد والپول، میربهادرقلی آریانژاد (مترجم)؛ دانشگاه علم و صنعت ایران
- ۵- «آمار و احتمالات کاربردی مهندسی». داگلاس سی. مونت گومری، کریم آتشگر (مترجم)؛ دانشگاه علم و صنعت ایران
- ۶- «اقتصادسنجی همراه با کاربرد EViews، Stata و R»، علی سوری، نشر نور علم



منابع لاتین

- 1- Washington, Simon P., Matthew G. Karlaftis, and Fred L. Mannering. **Statistical and econometric methods for transportation data analysis**. CRC press, 2010.
- 2- Spiegelman, Clifford, Eun Sug Park, and Laurence R. Rilett. **Transportation statistics and microsimulation**. CRC Press, 2010.
- 3- Ross, Sheldon M. **Introduction to probability models**. Academic press, 2014.
- 4- Wooldridge, Jeffrey. **Introductory econometrics: A modern approach**. Cengage Learning, 2012.



روش های تهیه کتاب؛

۱- تهران ، نارمک ، دانشگاه علم و صنعت ایران ، درب ورودی بانک ملت
مرکز فروش انتشارات دانشگاه علم و صنعت ایران
شنبه تا چهارشنبه ۸ الی ۱۳

۲- تهران ، تهرانپارس ، بالاتر از فلکه دوم
خیابان شهید نقلی اصفهانی ۱۸۴ شرقی ، پلاک ۶۹ واحد یک
انتشارات دیم

تلفن تماس: ۱۳ ۷۷۷۲۵۴ - ۰۲۱ داخلی ۷ و ۸
شنبه تا چهارشنبه ساعت ۹ تا ۱۷



دلایل ارائه این درس

- ۱- درک اندک دانشجویان کارشناسی ارشد و دکتری مهندسی حمل و نقل، از مفاهیم و اصطلاحات آماری و مدل‌های اقتصادسنجی
- ۲- مشکلات فراوان دانشجویان در هنگام مطالعه کتب و مقالات تخصصی، به دلیل ناآشنایی با اصطلاحات و مفاهیم پایه‌ای آماری و اقتصادسنجی
- ۳- مشاهده اشتباهات فراوان و اساسی در پایان نامه‌های دانشجویان



اهداف درس

- ۱- ارائه یک بستر مناسب برای آشنایی با اصول مبانی اولیه و اساسی استنباط آماری و مدلسازی آماری با تمرکز بر مباحث رایج در حوزه مهندسی حمل و نقل و ترافیک
- ۲- بررسی دقیق مفاهیم و اصطلاحات پایه‌ای آماری
- ۳- آشنایی با اصطلاحات لاتین برای تسهیل مطالعه کتب و مقالات تخصصی (در پیوست کتاب موجود است)



نمره درس

۱- آزمون ۱۲ نمره

۲- حل تمرین ۲ نمره

۳- پروژه درسی ۶ نمره

- مطالعه منابع، آشنایی با یک مدل و ارائه توضیحات مفصل به همراه رفرنس (۳ نمره)
- ساخت مدل با استفاده از داده‌های مشخص در محیط R (۱.۵ نمره)
- تفسیر نتایج (۰.۵ نمره)
- تحویل کد R و متن گزارش (۱ نمره)



1. مقدمه و آشنایی با انواع داده‌ها و متغیرها (و مفهوم متغیر تصادفی)
2. مباحث پایه احتمالات و کاربردهای آن در حوزه حمل‌ونقل
3. توزیع‌های آماری گسسته و پیوسته (با تاکید بر موارد پر کاربرد در حوزه حمل‌ونقل)
4. توزیع‌های نمونه‌ای
5. آزمون‌های آماری پارامتریک
6. آزمون‌های آماری ناپارامتریک
7. رگرسیون خطی ساده
8. موارد نقض فرض کلاسیک (آزمون‌ها شناسایی و راهکارهای جایگزین)
9. انواع مدل‌های رگرسیون داده‌های شمارشی (پواسون، دو جمله‌ای منفی، و...)
10. انواع مدل‌های رگرسیون لوجستیک، لوجیت، پروبیت
11. رویکرد بیزین

فصل اول



مقدمه ✓



تعریف علم آمار

گردآوری، توصیف، تجزیه و تحلیل و تفسیر داده‌های کمی و کیفی با استفاده از داده‌های جمع‌آوری شده از جامع آماری.

آمار توصیفی

✓ ارائه توصیف عددی و گرافیکی در مورد داده‌های جمع‌آوری شده از نمونه‌گیری

✓ ارائه استنباط‌های کیفی در مورد وضعیت پارامترهای جامعه

آمار استنباطی

✓ روش‌هایی که با استفاده از آنها و با در نظر گرفتن یک سطح اطمینان معین، می‌توان نتایج بدست آمده از نمونه تصادفی را به جامعه تعمیم داد.



تعریف اقتصادسنجی (Econometrics)

- اقتصادسنجی ترکیب شده از علم آمار، نظریه‌های اقتصادی و ریاضیات است
- به طور خلاصه، علم اقتصادسنجی در مورد انواع مدل‌های رگرسیون و نحوه برآورد آنها بحث می‌کند.
- در واقع، بسیاری از روش‌های اقتصادسنجی، کاربرد مدل‌های آماری را بیان می‌کنند.
- اما روش‌های اقتصادسنجی تنها برای مطالعه پدیده‌های اقتصادی استفاده نمی‌شود.



بخش‌های اصلی درس







□ جامعه آماری

مجموعه‌ای از افراد، اشیاء، پدیده‌ها، و یا ... که حداقل دارای یک خصوصیت مشترک باشند. مثلاً؛

- مصدومین تصادفات ترافیکی شهر مشهد
- رانندگان ناوگان اتوبوس‌رانی درون شهری تهران

□ نمونه

هر زیرمجموعه از جامعه، یک نمونه نامیده می‌شود.



تفاوت «احتمال» با «آمار»

در **احتمال**، فرض بر این است که تحلیل گر از توزیع پارامتر موردنظر در جامعه آگاه بوده و هدف او محاسبه احتمال مشاهده یک رویداد خاص است.

- اگر بدانیم تعداد تصادفات در یک بازه زمانی معین از یک توزیع پواسون با پارامتر (λ) ، پیروی می کند، آنگاه می توان احتمال مشاهده X_0 تصادف در یک بازه زمانی دلخواه را بدست آورد.

در **آمار**، مشخصات یا توزیع پارامترهای جامعه اغلب نامعلوم است و هدف، برآورد مشخصه مجهول جامعه براساس داده های نمونه می باشد.

- داده های "سرعت تردد خودروها" در قطعه ای از یک بزرگراه را به کمک نمونه گیری جمع آوری کرده باشیم و بخواهیم به کمک آن، "میانگین سرعت تردد خودروها" در آن قطعه را برآورد کنیم.



- **پارامتر.** مشخصه مجهول جامعه که قرار است در مورد آن استنباطی انجام شود.
برای مثال اگر جامعه را به صورت «زمان سفر وسایل نقلیه در کمان A» در نظر بگیریم. آنگاه، پارامتر جامعه می تواند «میانگین زمان سفر (μ) تمامی وسایل نقلیه عبوری از کمان مورد نظر» باشد.

- **آماره (برآوردگر).** هر تابعی که به کمک آن، برآورد انجام می گردد؛ برآوردگر نامیده می شود.
برای مثال، اگر یک نمونه از «زمان سفر وسایل نقلیه در کمان A» جمع آوری کنیم. آنگاه، میانگین نمونه (\bar{X}) یک برآوردگر برای میانگین جامعه (μ) است.



برخی از اصطلاحات و علائم رایج در آمار و احتمالات

نسبت	اندازه	ضریب همبستگی	انحراف معیار	واریانس	میانگین	مشخصه		
\hat{p}	n	r	S	S^2	\bar{X}	آماره (برآوردگر)	نمونه	آمار توصیفی
p	N	ρ	σ	σ^2	μ	پارامتر	جامعه	آمار استنباطی



□ اندازه‌گیری (measurement)

- **مقیاس اندازه‌گیری:** یکی از مشخصات داده‌ها (مشاهدات) که قبل از هر گونه تجزیه و تحلیل آماری لازم است به آن توجه کنیم.
- **اندازه‌گیری** عبارت است از نسبت دادن اعداد به خصوصیات اشیاء، رخدادها، یا افراد.
- **مقیاس‌های اندازه‌گیری** به چهار دسته اسمی، رتبه‌ای، فاصله‌ای و نسبتی تقسیم می‌شوند.



□ مقیاس‌های اندازه‌گیری (measurement Scales)

❖ **مقیاس اسمی (nominal scale)** تنها برای نامگذاری به کار می‌رود. اطلاعات با مقیاس

اسمی را نمی‌توان به صورت معنی‌دار از کم به زیاد مرتب نمود.

✓ در اینجا، اعداد صرفاً یک برچسب هستند و نمی‌توان بر روی آن‌ها عملیات ریاضی انجام داد.

❖ مثال: محل تولد، رنگ، جنسیت، گروه خونی، وضعیت تاهل



□ مقیاس‌های اندازه‌گیری (measurement Scales)

❖ **مقیاس رتبه‌ای یا ترتیبی (ordinal scale)** هرگاه مقیاس یک متغیر که یک عدد است، برتری و ترتیب را بیان کند.

✓ صرفاً یک برچسب نیستند و یک ترتیب ذاتی درون خود دارند. اما باز هم نمی‌توان بر روی آنها عملیات ریاضی انجام داد.

❖ مثال: کیفیت روسازی (عالی، خوب، متوسط، ضعیف)، مدرک تحصیلی (زیردیپلم، دیپلم، فوق دیپلم، لیسانس و بالاتر)



□ مقیاس‌های اندازه‌گیری (measurement Scales)

- ❖ **مقیاس فاصله‌ای (Interval scale)** هرگاه مقیاس یک متغیر که یک عدد حقیقی است، فاصله بین مقادیر را حفظ می‌کند، اما نسبت را حفظ نمی‌کند.
- ❖ شبیه به مقیاس نسبتی است با این تفاوت که **صفر مطلق ندارد**. به این معنا که وقتی مقدار متغیر برابر با صفر است، به این معنا نیست که هیچ مقداری از آن متغیر وجود ندارد.
- ❖ نمونه: درجه حرارت (برای مثال نمی‌توان گفت که درجه حرارت ۱۰ درجه سلسیوس، دقیقا ۲ برابر گرم‌تر از درجه حرارت ۵ درجه سلسیوس است؛ زیرا ۱۰ درجه سلسیوس معادل ۵۰ درجه فارنهایت است و ۵ درجه سلسیوس معادل ۴۱ درجه فارنهایت)
- ❖ تعداد متغیرهای با مقیاس فاصله‌ای بسیار اندک است (مثلا نمرات آزمون‌های رفتار رانندگی) و در تحلیل‌های آماری **تقریبا هیچ تفاوتی بین متغیرهای فاصله‌ای و نسبتی وجود ندارد**.



□ مقیاس‌های اندازه‌گیری (measurement Scales)

- ❖ **مقیاس نسبتی (ratio scale)** هرگاه مقیاس یک متغیر که یک عدد حقیقی است، نسبت را حفظ کند؛ به آن مقیاس نسبتی گویند. برای مثال، اگر وزن A و B با مقیاس گرم برابر با 6000 گرم و 2000 گرم باشد، با مقیاس کیلوگرم، برابر با 6 و 2 خواهد بود. با هر دو مقیاس، وزن A ، 3 برابر B است.
- ✓ فاصله بین مقادیر، معنادار است.
- ✓ نسبت بین مقادیر معنادار است.
- ❖ مثال: عرض راه، سرعت، مدت زمان سبز، سن، زمان عکس‌العمل



□ مقیاس‌های اندازه‌گیری (measurement Scales)

- ❖ **مقیاس شمارشی (Counting scale)** شبیه به مقیاس نسبتی است با این تفاوت که عدد حقیقی نیست بلکه عدد طبیعی می‌باشد.
- ✓ چهار عمل اصلی ریاضی در مورد آن‌ها صدق می‌کند (در بازه اعداد طبیعی).
- ❖ مثال: تعداد تصادفات، تعداد خطوط یک خیابان (lane)



□ متغیر

نمادی است که مقادیر یا ارزش‌هایی به آن منتسب می‌شود.

□ انواع متغیرها

❖ متغیرهای کیفی (qualitative): متغیرهایی که (واحد اندازه‌گیری ندارد) با مقیاس‌های اسمی یا رتبه‌ای اندازه‌گیری می‌شوند.

➤ مانند: نوع سفر (کاری، تفریحی، تحصیلی، خرید)، جنسیت راننده (مذکر، مونث)، مدرک تحصیلی

❖ متغیرهای کمی (Quantitative): متغیرهایی که از طریق شمارش یا با مقیاس‌های فاصله‌ای یا نسبتی اندازه‌گیری می‌شوند.

➤ مانند: تعداد سفرهای کاری تولید شده از یک منطقه، سرعت تردد خودروها، بُعد خانوار، تعداد تصادفات
سالیانه در یک شهر



□ انواع داده‌ها

❖ **داده‌های گسسته (discrete):** داده‌هایی که از طریق اندازه‌گیری با مقیاس‌های اسمی، رتبه‌ای یا شمارشی به دست می‌آیند.

➤ مانند: تعداد سفر (کاری، تفریحی، تحصیلی، خرید)، تعداد راننده (مذکر، مونث)، تعداد افراد دارای مدرک تحصیلی کارشناسی

❖ **داده‌های پیوسته (Continuous):** داده‌هایی که از طریق اندازه‌گیری با مقیاس‌های فاصله‌ای یا نسبتی به دست می‌آیند.

➤ مانند: سرعت تردد خودروها، عرض راه