



استنباط آماری (پیش‌درآمدی بر علوم داده و یادگیری ماشین (۱))

نویسندگان

دیمیتری پی. برتسکاس

جان ال. توستیکلیس

ترجمه و تخلص

دکتر احمد خونساری

(عضو هیئت علمی دانشکده مهندسی برق و کامپیوتر دانشگاه تهران)

۱۳۹۹

اداره انتشارات دانشگاه شاهد (۱۸۹)

سرشناسه	: برتسکاس، دیمیتری پی. Bertsekas, Dimitri P
عنوان و نام پدیدآور	: استنباط آماری (پیش‌درآمدی بر علوم داده و یادگیری ماشین ۱) / نویسندگان: دیمیتری پی. برتسکاس، جان آن. تسیتسیکلیس؛ ترجمه و تخلص: احمد خونساری.
مشخصات نشر	: تهران: دانشگاه شاهد، معاونت پژوهش و فناوری، مرکز چاپ و انتشارات، ۱۳۹۹.
مشخصات ظاهری	: ۲۹۸ ص.: نمودار.
فروست	: اداره انتشارات دانشگاه شاهد؛ ۱۸۹.
شابک	: ۹۷۸-۶۰۰-۶۱۲۱-۸۶-۴
وضعیت فهرست نویسی	: فیبا
موضوع	: احتمالات
موضوع	: Probabilities
موضوع	: آمار ریاضی
موضوع	: Mathematical statistics
شناسه افزوده	: تسیتسیکلیس، جان آن.
شناسه افزوده	: Tsitsiklis, John N.
شناسه افزوده	: خونساری، احمد، ۱۳۴۱-، مترجم
شناسه افزوده	: دانشگاه شاهد، مرکز چاپ و انتشارات
رده‌بندی کنگره	: QA۲۷۳
رده‌بندی دیویی	: ۵۱۹/۲
شماره کتابشناسی ملی	: ۷۴۱۶۸۷۲
وضعیت رکورد	: فیبا



کتابخانه
دانشگاه گیلان
معاونت پژوهش و فناوری
اداره انتشارات

عنوان کتاب: استنباط آماری (پیش‌درآمدی بر علوم داده و یادگیری ماشین ۱)

ترجمه و تخلص: احمد خونساری

تعداد: ۵۰۰ نسخه

نوبت و سال چاپ: اول - ۱۳۹۹

قیمت: ۶۷۷۰۰۰ ریال

نشانی ناشر: تهران، ابتدای آزادراه خلیج فارس، روبه‌روی حرم مطهر امام خمینی (ره)، معاونت پژوهش و

فناوری، اداره انتشارات دانشگاه شاهد

دورنگار: ۵۱۲۱۵۱۲۴

تلفن: ۵۱۲۱۴۱۲۲

فهرست مطالب

۵	مقدمه
ط	فهرست اشکال
۱	فصل اول احتمال
۲	۱.۱ مروری بر احتمال
۲	۱.۱.۱ متغیرهای تصادفی
۲	۲.۱ مجموعه‌های شمارا
۴	۳.۱ متغیرهای تصادفی گسسته
۷	۴.۱ امید
۱۰	۵.۱ پیشامدها و احتمالات
۱۱	۶.۱ استقلال و وابستگی
۱۲	۷.۱ احتمال شرطی
۲۰	۸.۱ مسائل
۳۱	فصل دوم مباحث تکمیلی در متغیرهای تصادفی
۳۱	۱.۲ کواریانس و همبستگی
۳۶	۲.۲ واریانس مجموع چند متغیر تصادفی

۳۹	امید شرطی و واریانس شرطی	۳.۲
۴۳	امید شرطی به‌عنوان یک تخمین‌گر	۴.۲
۴۵	واریانس شرطی	۵.۲
۵۰	مسائل	۶.۲

۶۱ فصل سوم قضایای حدی

۶۳	نا برابری‌های مارکوف و چبی‌شف	۱.۳
۶۹	قانون ضعیف اعداد بزرگ	۲.۳
۷۲	همگرایی در احتمال	۳.۳
۷۶	قضیه حد مرکزی	۴.۳
۷۸	تقریب‌های مبتنی بر قضیه حد مرکزی	۱.۴.۳
۸۵	قانون قوی اعداد بزرگ	۵.۳
۸۷	همگرایی با احتمال یک	۱.۵.۳
۸۹	خلاصه و بحث	۶.۳
۹۰	مسائل	۷.۳

۱۰۷ فصل چهارم استنباط آماری بیزی

۱۱۴	استنباط بیزی و توزیع پسین	۱.۴
۱۲۵	تخمین‌زدن نقطه‌ای، آزمون فرض و قاعده MAP	۲.۴
۱۴۰	تخمین‌زدن حداقل میانگین مربعات بیزی	۳.۴
۱۵۰	تخمین‌زدن "حداقل میانگین مربعات" خطی بیزی	۴.۴
۱۶۰	خلاصه	۵.۴
۱۶۰	مسائل	۶.۴

مقدمه مترجم

هم‌زمان با افزایش نرخ تولید داده و به دنبال آن ازدیاد تنوع و حجم داده، شاهد رشد بیش از پیش علم داده هستیم. از این رو، کسب مهارت در زمینه‌ی علوم آماری، جهت مدیریت و تحلیل داده، امری لازم است. امروزه حوزه‌ی علمی جدیدی موسوم به آمار و علوم داده ایجاد شده است که در آن به کمک علوم آمار، داده‌ها تجزیه و تحلیل می‌شوند. به این ترتیب این حوزه جدید بنیانی متقن برای تفسیر نتایج و استخراج مفاهیم از داده‌ها ایجاد می‌کند. همان‌گونه که از تعریف علم داده مشهود است یکی از ابزارهای زیربنایی آن علوم آماری است. در واقع علوم آماری چارچوب نظری و همچنین ابزار عملی جهت تحقق اهداف تعریف شده در علم داده را فراهم می‌آورند. فراگیری علوم آماری و علم داده و به کارگیری آنها منجر به درک، تجزیه و تحلیل، تفسیر و مشاهده دقیق پدیده‌های واقعی خواهد شد.

به‌طور کلی مطالعه علوم آماری و علم داده باعث ایجاد و ارتقاء توانایی‌های زیر می‌شود:

- مهارت‌های ارتباطی، آماری و محاسباتی
- مدیریت حجم زیاد اطلاعات
- به اشتراک‌گذاری یافته‌ها و دیدگاه‌ها در حوزه‌ی داده با متخصصان دیگر و یا مخاطب عام
- ساخت مدل‌های ریاضی مبتنی بر داده
- مقایسه آماری مدل‌های ریاضی متنوع و انتخاب مدل برتر
- ایجاد چشم‌اندازهای حرفه‌ای در نهادهای دولتی و خصوصی

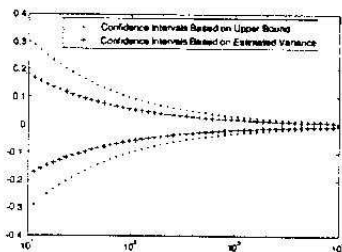
- امکان کسب مهارت‌های تخصصی جهت پژوهش‌های آکادمیک

این کتاب ترجمه‌ای گزینشی از چاپ دوم کتاب «مقدمه‌ای بر احتمالات» نوشته دیمیتری برتسکاس و جان تسیتسیکلیس اساتید مؤسسه فناوری ماساچوست (MIT) می‌باشد. در این کتاب نظریه‌های آماری دقیق با تکیه بر تجربه‌های عملی، که می‌تواند به عنوان الگویی جهت به کارگیری مدل‌های آماری در تجزیه و تحلیل داده‌ها باشد، بیان شده است. به بیانی دیگر در این کتاب به جای تکیه بر تکنیک‌ها، بر نظریه‌ها و مفاهیم پایه آماری تأکید شده است.

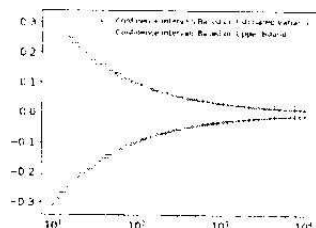
متن حاضر برای طیف وسیعی از افراد با سطوح دانش متفاوت که علاقه‌مند به استفاده از روش‌های آماری مدرن جهت مدل‌سازی، تحلیل، پیش‌بینی و استخراج نتایج جدید از داده‌ها هستند، در نظر گرفته شده است. گروه هدف این کتاب می‌تواند دانشجویان، پژوهشگران، مهندسیین و به‌طور خاص تحلیلگران داده باشند. همچنین با توجه به کاربرد وسیع علوم آماری در حوزه‌های مختلف، مطالب این کتاب می‌تواند برای متخصصین و دانشجویان شاغل در حوزه‌های پزشکی، زیست‌شناسی، زمین‌شناسی، روان‌شناسی، علوم اجتماعی و کسب و کار نیز مفید واقع شود. برای مطالعه این کتاب دانستن مفاهیم اولیه احتمال مورد نیاز می‌باشد. از این رو، فصل اول این کتاب به صورت موجز به مفاهیم کلیدی احتمال پرداخته است تا نیاز خواننده از این منظر نیز پوشش داده شده باشد.

از تلاش‌های آقای مهدی دولتی، آقای دکتر ابوالفضل دیانت و آقای محمد برهانی در جهت آماده‌سازی کتاب تشکر می‌شود.

خطاها و تفاوت‌ها: طی ترجمه این کتاب سعی شده است که حداکثر امانت رعایت شود و محتوا بدون اشکال به خواننده ارائه گردد. اما مشاهده خطا و مشکل حادثه دور از انتظار نیست. از خوانندگان دقیق و باملاحظه تقاضا می‌شود در صورت برخورد به خطاهای سهوی نگارشی آن‌ها را از طریق بستر الکترونیکی با مؤلف در میان بگذارند. همچنین در ادامه به ذکر یک مورد عدم همخوانی نمودار و معادلات ریاضی در کتاب اصلی می‌پردازیم و نتایج خود را بیان می‌کنیم. در فصل پنجم کتاب حاضر، ذیل مثال ۸.۵ دو شیوه محاسبه بازه اطمینان در شکل ۴.۵ مقایسه شده‌اند. نمودارهای کتاب مرجع با معادلات داخل مثال



شکل مندرج در کتاب اصلی (ب)



(آ) پیاده‌سازی مترجم

شکل ۱: تمایز نمودار و معادلات ریاضی در کتاب اصلی و نتایج مؤلف

همخوانی ندارند. در شکل ۱۱ حاصل پیاده‌سازی‌های مترجم و در شکل ۱۲ شکل اصلی کتاب را ملاحظه می‌کنید.

مقدمه مؤلفین کتاب

این کتاب بر اساس مطالب درس «تحلیل سیستم‌های احتمالی» در مؤسسه فناوری ماساچوست است. در درس مذکور طیف وسیعی از دانشجویان با زمینه‌ها و علایق متنوع حضور داشته‌اند. این دانشجویان شامل مقاطع کارشناسی و تحصیلات تکمیلی از دانشکده‌های مهندسی تا مدیریت بوده‌اند. بر این اساس، ما سعی کرده‌ایم بین سادگی در ارائه مطالب و پیچیدگی استدلال‌های تحلیلی موازنه‌ای ایجاد نماییم.

هدف اصلی ما گسترش مهارت‌هایی است که بتوان توسط آن مدل‌های احتمالی را، به شیوه‌ای که درک بصری و دقت ریاضی را ترکیب می‌کند، ساخته و مورد استفاده قرار دهیم. به این ترتیب در متن به برخی از تجزیه و تحلیل‌های ریاضی فقط اشاره کرده‌ایم و یا به صورت شهودی توضیح داده‌ایم. بنابراین اثبات‌های پیچیده بیان نشده‌اند و در مواردی نیز با بیانی ساده آنها را ذکر کرده‌ایم. در عین حال، برخی از این تحلیل‌ها و نتایج مورد نیاز ریاضی (در سطح حساب پیشرفته) در مسائل نظری در پایان فصل مربوطه گنجانده شده است. علاوه بر این، به برخی از مسائل ریاضی ظریف نیز در پانویس‌ها اشاره شده است.

پس از فصل ۱ که احتمال را به طور اجمالی مرور می‌کند، در فصل ۲ مفاهیم وابستگی متغیرهای

تصادفی و امید شرطی بیان می‌شود که پیش‌نیاز استنباط آماری است.