

به نام یزدان پاک

همه سوال های

دکترای تخصصی (Ph.D)

آمار

(جلد ۱)

دانشگاه آزاد

۱۳۷۹-۱۳۸۸

تالیف و تدوین

احمد ملک زاده

پریسا یوسفی زوج

همه سوال‌های دکترای تخصصی (Ph.D) آمار جلد ۱ (دانشگاه آزاد)

تالیف و تدوین	: احد ملک‌زاده، پریسا یوسفی زوج
ناشر	: پردازش
چاپ دوم	: ۱۴۰۲
تعداد	: ۱۰۰
حروفچینی	: پردازش
لیتوگرافی	: پردازش
چاپ و صحافی	: پردازش
قیمت	: ۴۰۰.۰۰۰ تومان

حق چاپ محفوظ و متعلق به نشر پردازش است.

سرشناسه	: ملک‌زاده، احد
عنوان و پدیدآور	: همه سوال‌های دکترای تخصصی آمار (دانشگاه آزاد)
مشخصات نشر	: تالیف و تدوین: احد ملک‌زاده، پریسا یوسفی زوج تهران: پردازش، ۱۴۰۲.
مشخصات ظاهری	: ۱۷۰ ص: ۲۹×۲۲ س.م.
شابک	: ۹۷۸-۹۶۴-۲۰۳-۹۷۰-۸
وضعیت فهرست‌نویسی	: فیپا
موضوع	: آمار -- راهنمای آموزشی (عالی)
موضوع	: دانشگاه‌ها و مدارس عالی - ایران - آزمون‌ها - راهنمای مطالعه
موضوع	: آزمون دوره‌های تحصیلات تکمیلی - ایران
شناسه افزوده	: پریسا یوسفی زوج
رده‌بندی کنگره	: ۱۳۷۸ خ ۸۴ ج / ۳۰۵ JZ
رده‌بندی دیویی	: ۳۲۷/۱۰۱
شماره کتابخانه ملی	: ۱۵۸۶۴۴۵

نشر پردازش

تهران: میدان انقلاب، خیابان فروردین، خیابان شهدای ژاندارمری، روبه‌روی اداره پست، پلاک ۱۳۲ و ۱۳۰

ساختمان پردازش

تلفن: ۶۶۹۷۵۵۶۶ (۱۰۰ خط) دورنگار: ۶۶۴۰۷۴۰۸

www.Pardazesh.org

www.Pardazeshpub.com

سایت برتر کشور در چهارمین و پنجمین همایش ملی تجارت الکترونیک

فهرست مطالب

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۴

۸۰	آنالیز ریاضی.....
۸۴	احتمال پیشرفته.....
۸۷	استنباط آماری.....
۹۲	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۵

۹۴	آنالیز ریاضی.....
۹۶	احتمال پیشرفته.....
۹۹	استنباط آماری.....
۱۰۴	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۶

۱۱۰	آنالیز ریاضی.....
۱۱۳	احتمال پیشرفته.....
۱۱۶	استنباط آماری.....
۱۲۴	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۷

۱۲۸	آنالیز ریاضی.....
۱۳۱	احتمال پیشرفته.....
۱۳۶	استنباط آماری.....
۱۴۲	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۸

۱۵۰	آنالیز ریاضی.....
۱۵۳	احتمال پیشرفته.....
۱۵۶	استنباط آماری.....
۱۶۳	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۷۹

۵	آنالیز ریاضی.....
۸	احتمال پیشرفته.....
۱۱	استنباط آماری.....
۱۸	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۰

۲۰	آنالیز ریاضی.....
۲۳	احتمال پیشرفته.....
۲۷	استنباط آماری.....
۳۲	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۱

۳۴	آنالیز ریاضی.....
۳۷	احتمال پیشرفته.....
۴۰	استنباط آماری.....
۴۴	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۲

۴۶	آنالیز ریاضی.....
۴۹	احتمال پیشرفته.....
۵۳	استنباط آماری.....
۵۸	زبان تخصصی انگلیسی.....

آزمون دکترای تخصصی آمار «دانشگاه آزاد»

۱۳۸۳

۶۰	آنالیز ریاضی.....
۶۶	احتمال پیشرفته.....
۷۰	استنباط آماری.....
۷۸	زبان تخصصی انگلیسی.....

پیشگفتار

دانشگاه آزاد اسلامی با سابقه درخشان و طولانی در امر آموزش عالی، در تمامی مقاطع و رشته‌های تحصیلی از دیرباز مشغول به فعالیت بوده است. رشته آمار به عنوان یکی از رشته‌های آموزشی این دانشگاه در مقاطع تحصیلی کارشناسی و کارشناسی ارشد مورد توجه بوده و در مقطع دکتری از سال ۷۹ به بعد با برگزاری آزمون ورودی شروع به جذب دانشجو کرده است. این آزمون‌ها به صورت منظم همه‌ساله در بهمن‌ماه برگزار می‌گردد. دانشجویان و فارغ‌التحصیلان کارشناسی ارشد علاقه‌مند به این رشته می‌توانند با شرکت در این آزمون و پس از کسب نمره قبولی در آزمون و مصاحبه مشغول به تحصیل شوند.

در این آزمون از دروس استنباط آماری ۱ و ۲، احتمال پیشرفته، زبان تخصصی و آنالیز ریاضی سوالاتی به صورت مجزا طراحی می‌شود. آزمون دروس آنالیز ریاضی و احتمال پیشرفته صبح و آزمون دروس استنباط آماری و زبان تخصصی بعدازظهر همان روز برگزار می‌گردند. مدت در زمان (برحسب دقیقه) آزمون‌های برگزار شده و تعداد سوالات در جدول زیر به صورت خلاصه آورده شده‌اند.

سال	آنالیز ریاضی		احتمال پیشرفته		استنباط آماری		زبان
	تعداد	زمان	تعداد	زمان	تعداد	زمان	زمان
۷۹	۶	۱۲۰	۶	۹۰	۸	۱۲۰	۴۵
۸۰	۵	۱۰۰	۵	۱۰۰	۶	۱۲۰	۴۵
۸۱	۴	۱۲۰	۵	۱۲۰	۴	۱۲۰	۵۰
۸۲	۷	۱۲۰	۷	۱۲۰	۶	۱۲۰	۶۰
۸۳	۸	۱۲۰	۷	۱۲۰	۶	۱۲۰	۳۰
۸۴	۵	۱۲۰	۶	۱۲۰	۴	۹۰	۶۰
۸۵	۳	۹۰	۳	۹۰	۳	۱۲۰	۷۰
۸۶	۵	۱۲۰	۵	۱۲۰	۴	۱۲۰	۶۰
۸۷	۴	۱۲۰	۶	۱۲۰	۳	۱۲۰	۱۲۰
۸۸	۴	۱۲۰	۵	۱۲۰	۳	۱۲۰	۱۲۰

نتایج حاصل نشان‌دهنده مدت زمان تقریباً ثابت ۱۲۰ دقیقه در امتحان‌های سال‌های اخیر دانشگاه آزاد اسلامی و کاهش تعداد سوالات، نسبت به سال‌های ابتدایی می‌باشد.

این کتاب به منظور جمع‌آوری مجموعه سوالات آزمون دکتری دانشگاه آزاد اسلامی در رشته آمار، همراه با پاسخ آنها تهیه و در اختیار علاقه‌مندان به این رشته قرار گرفته است. برخی از سوالات دارای یک راه‌حل خاص نیست، بلکه به روش‌های مختلف می‌توان آنها را پاسخ داد. در این مجموعه سعی شده است که همه راه‌های پاسخ به سوالات آورده شود. از جمله مشکلات در مسیر پاسخ‌گویی، ایرادات تاییبی احتمالی در بعضی سوالات بوده است که در صورت وجود در ابتدای پاسخ‌گویی به آن اشاره و صورت صحیح آن ذکر شده است.

مجموعه سوالات گردآوری شده در این کتاب به صورت سالانه و همانند ترتیبی است که در برگزاری این امتحان رعایت می‌شود تا با بهره‌گیری از این کتاب بتوانید در فضایی نزدیک به زمان برگزاری قرار بگیرید. با مطالعه پیوسته این کتاب به نحوه برگزاری امتحان و روندهای موجود در نوع سوالات، منابع امتحانی و چگونگی سوالات طراحی شده می‌توانید بیشتر آشنا شوید. به طور حتم این مجموعه خالی از اشکال نیست، بنابراین از خوانندگان گرامی خواهشمندیم با ارسال نقطه نظرات خود ما را در بهینه‌سازی هرچه بیشتر آن یاری نمایند.

در اینجا لازم است از تمامی دوستانی که در تهیه این مجموعه ما را یاری و راهنمایی کرده‌اند کمال تشکر را داشته باشیم.

به امید آنکه این مجموعه بتواند در پیشرفت شما نقشی داشته باشد.

احمد ملک‌زاده

آزمون دکترای تخصصی آمار «دانشگاه آزاد» ۱۳۷۹

آنالیز ریاضی



۱. ثابت کنید به ازای هر عدد حقیقی a ، یک و تنها یک عدد صحیح مانند n هست که:

$$n^r + n \leq a < (n+1)^r + (n+1)$$

پاسخ

برای هر عدد صحیح و نامنفی a تعریف می‌کنیم $n = \lfloor \sqrt[r]{a} \rfloor$.

حال اگر $a \geq n^r + n$ باشد آنگاه $n^r + n \leq a < (n+1)^r + (n+1)$ خواهد بود.

حالت دوم، زمانی است $a < n^r + n$. در این حالت $n^r \leq a < (n+1)^r + (n+1)$ و با قراردادن $n-1 = m$ داریم:

$$m^r + m \leq a < (m+1)^r + (m+1)$$

و حکم ثابت می‌شود.

۲. فرض کنید f تابعی است حقیقی بر بازه باز (a, b) و دارای مشتق f' باشد، به علاوه فرض کنید تابع f' بر (a, b) کراندار

باشد. ثابت کنید تابع f را می‌توان به تابعی پیوسته بر بازه بسته $[a, b]$ توسعه داد.

پاسخ

به دلیل اینکه تابع f بر بازه (a, b) مشتق دارد، بنابراین در این بازه پیوسته نیز می‌باشد برای هر $x \in (a, b)$ داریم $|f(x)| \leq m$

یعنی $\lim_{x \rightarrow a^+} f(x)$ و $\lim_{x \rightarrow b^-} f(x)$ نیز وجود دارد. با در نظر گرفتن این مطالب و به منظور توسعه دادن f می‌نویسیم

$$f^*(x) = \begin{cases} f(x) & a < x < b \\ \lim_{x \rightarrow a^+} f(x) & x = a \\ \lim_{x \rightarrow b^-} f(x) & x = b \end{cases}$$

که تابعی پیوسته بر بازه $[a, b]$ می‌باشد.

۳. ثابت کنید تابع حقیقی $f(x) = \left(1 + \frac{h}{x}\right)^x$

به ازای هر $h > 0$ بر بازه $(1, \infty)$ صعودی است و به ازای هر $h < 0$ بر بازه $(1, \infty)$ نزولی است.

پاسخ

تابع $\ln$ تابعی یک به یک است. بنابراین با $\ln$ گرفتن از طرفین ضابطه تعریف می‌کنیم.

$$g(x) = \ln f(x) = x \ln \left(1 + \frac{h}{x}\right)$$

و با مشتق گرفتن از تابع $g(x)$ ضابطه‌ای به فرم زیر به دست می‌آید.

$$g'(x) = \ln\left(1 + \frac{h}{x}\right) - \frac{h}{x+h} = \ln\left(1 + \frac{h}{x}\right) - \left(1 - \frac{x}{x+h}\right)$$

به ازای هر $h > 0$ تابع $g'(x)$ تابعی نزولی از x می‌باشد، زیرا $g''(x) = \frac{h}{x(x+h)^2} = \left(\frac{1}{h+x} - \frac{1}{x}\right)$ همیشه منفی است. بنابراین کمترین مقدار را به ازای هر $h > 0$ ثابت، زمانی می‌پذیرد که $x \rightarrow \infty$ میل کند.

$$\lim_{x \rightarrow \infty} g'(x) = \lim_{x \rightarrow \infty} \ln\left(1 + \frac{h}{x}\right) - \lim_{x \rightarrow \infty} \frac{h}{h+x} = 0$$

یعنی برای $h > 0$ ثابت، تابع $g'(x)$ همیشه مقدار مثبت می‌پذیرد و این بدین معناست که تابع اولیه صعودی است.

ب) اگر $h < 0$ باشد نیز تابع $f(x)$ تابعی صعودی است، زیرا $g'(x) = \ln\left(1 + \frac{h}{x}\right) - \frac{h}{x+h}$ به ازای $x > -h$ همیشه مثبت است.

۴. تابع α بر بازه $[0, 4]$ چنین تعریف می‌شود.

$$\alpha(x) = \begin{cases} x^2 & 0 \leq x \leq 2 \\ [x] - 2 & 2 < x \leq 4 \end{cases}$$

($[x]$ تابع جزء صحیح x است) ثابت کنید $\int_0^4 x^2 dx$ موجود است و مقدار آن را محاسبه کنید.

پاسخ

تابع $f(x) = x^2$ در تمام بازه $[0, 4]$ پیوسته می‌باشد، حال بازه $[0, 4]$ را به دو بازه $[0, 2]$ و $[2, 4]$ تفکیک می‌کنیم $\alpha(x)$ نیز در این بازه‌ها تابعی صعودی است. بنابراین با تفکیک انتگرال به سه انتگرال مجزا به صورت زیر:

$$\int_0^4 x^2 dx = \int_0^2 x^2 dx + \int_2^3 x^2 dx + \int_3^4 x^2 dx$$

هر سه انتگرال موجود و قابل محاسبه می‌باشند، یعنی انتگرال کل وجود دارد. با محاسبه انتگرال‌ها و جمع مقادیر حاصله داریم:

$$\int_0^4 x^2 dx = \frac{2}{3} x^3 \Big|_0^2 + 1 \times 9 + 1 \times 16 = 33$$

۵. فرض کنیم f تابعی حقیقی و مشتق‌پذیر بر بازه باز (a, b) بوده، به ازای یک $c \in (a, b)$ داشته باشیم $f(c) = 0$. $f'(c) \neq 0$ به علاوه برای هر $f(x) \neq 0, x \neq c$.

ثابت کنید $\int_a^b \frac{dx}{f(x)}$ موجود نیست.

پاسخ

بنا به فرض سوال $f(c) = 0$ و $f'(c) \neq 0$ بنا به این فرضیات دو حالت برای $f'(c)$ پیش می‌آید. در حالت اول می‌توان فرض کرد $f'(c) = x > 0$. بنابراین در سمت چپ نزدیک c ، تابع f مقادیر منفی و در سمت راست نزدیک c ، مقادیر مثبت می‌پذیرد. حال بنا به فرض تابع f پیوسته و در جای دیگر غیر از c مقدار آن صفر نمی‌شود. این بدان معناست که تابع f در بازه (c, b) مثبت و در بازه (a, c) مقدار آن منفی می‌باشد. حال تابع $\frac{1}{f}$ را در نظر بگیریم. این تابع در همه جا به جزء c پیوسته است و

$$\lim_{x \rightarrow c^+} \frac{1}{f} = +\infty \quad \lim_{x \rightarrow c^-} \frac{1}{f} = -\infty$$



بنابراین تابع $\frac{1}{f}$ در C تعریف نشده و در نزدیک آن بی کران است، یعنی تابع $\frac{1}{f}$ دارای انتگرال روی بازه (a, b) نمی باشد. در حالت دوم می توان فرض کرد $\alpha = f'(c) < 0$ ، در این حالت مطالب فوق الذکر عکس می شوند، ولی بنا به تعریف تابع $\frac{1}{f}$ ، دوباره این تابع در تمام بازه به جز c پیوسته می شود، یعنی تابع $\frac{1}{f}$ دارای انتگرال روی بازه (a, b) نمی باشند.

۶. دنباله توابع $\{f_n\}$ ، $\{g_n\}$ بر R چنین تعریف می شود.

$$f_n(x) = \frac{x^n}{1+x^{2n}} \quad \text{و} \quad g_n(x) = (x-1)f_n(x)$$

ثابت کنید دنباله توابع $\{f_n\}$ بر $(1, \infty)$ همگرایی یکنواخت نیست، ولی دنباله توابع $\{g_n\}$ بر $(1, \infty)$ همگرایی یکنواخت است.

پاسخ

مجموعه L را تمامی دنباله های بزرگ تر از یک و همگرا به نقطه ۱ در نظر بگیرید.

$$L = \left\{ \{x_n\} \mid \forall_n x_n > 1 \text{ \& } \lim_{n \rightarrow \infty} x_n = 1 \right\}$$

آنگاه برای دنباله $\{g_n\}$ عضو L داریم:

$$\lim_{n \rightarrow \infty} f_n(g_n) = \lim_{n \rightarrow \infty} f_n(1) = \frac{1}{2}$$

در حالی که تابع f_n برای هر $x > 1$ زمانی که n به بی نهایت میل می کند مقداری برابر صفر می پذیرد، یعنی:

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \frac{x^n}{1+x^{2n}} = 0 \quad x > 1$$

بنابراین $f_n(x)$ همگرایی یکنواخت نیست.

برای تابع $g_n(x) = (x-1)f_n(x)$ نیز

$$\lim_{n \rightarrow \infty} g_n(x) = \lim_{n \rightarrow \infty} (x-1)f_n(x) = \lim_{n \rightarrow \infty} \frac{(x-1)x^n}{1+x^{2n}} = 0 \quad \text{for } x > 1$$

حال اگر دنباله $\{z_n\}$ عضو L باشد، آنگاه

$$\lim_{n \rightarrow \infty} g_n(z_n) = \lim_{n \rightarrow \infty} (z_n - 1)f_n(z_n) = 0$$



احتمال پیشرفته

۷. فرض کنید $\lim_{n \rightarrow +\infty} p(A_n) = 1$ و $\lim_{n \rightarrow +\infty} p(B_n) = p$ ثابت کنید:

$$\lim_{n \rightarrow +\infty} p(A_n \cap B_n) = p$$

پاسخ

برای هر $n \in \mathbb{N}$ داریم که $A_n \cap B_n \subset B_n$ بنابراین:

$$p(A_n \cap B_n) \leq p(B_n)$$

و براساس نامساوی بونفروی برای هر $n \in \mathbb{N}$

$$p(A_n \cap B_n) \geq p(A_n) + p(B_n) - 1$$

بنابراین:

$$p(A_n) + p(B_n) - 1 \leq p(A_n \cap B_n) \leq p(B_n)$$

حال اگر n را به سمت بی نهایت میل دهیم نتیجه به فرم زیر می شود

$$\lim_{n \rightarrow \infty} p(A_n) + \lim_{n \rightarrow \infty} p(B_n) - 1 \leq \lim_{n \rightarrow \infty} p(A_n \cap B_n) \leq \lim_{n \rightarrow \infty} p(B_n)$$

یعنی:

$$p \leq \lim_{n \rightarrow \infty} p(A_n \cap B_n) \leq p$$

بنا به قضیه فشردگی داریم

$$\lim_{n \rightarrow \infty} p(A_n \cap B_n) = p$$

۸. متغیر تصادفی را تعریف کنید و ثابت کنید مجموع دو متغیر تصادفی نیز یک متغیر تصادفی است.

پاسخ

اگر (Ω, F, P) فضای اندازه باشد و تابع $f: \Omega \rightarrow \mathbb{R}$ اندازه پذیر، آنگاه f را یک متغیر تصادفی می گویند.

منظور از اندازه پذیر بودن f این است که برای هر عضو فضای بورل همانند B داشته باشیم $f^{-1}(B) \in F$.

ب) برای اثبات اینکه تابعی همانند f یک متغیر تصادفی می باشد کافی است ثابت کنیم که برای هر $c \in \mathbb{Q}$ (مجموعه اعداد گویا)

$$\{w, f(w) \leq c\} \in F$$

حال فرض می کنیم f و g دو متغیر تصادفی روی فضای (Ω, F, P) باشند آنگاه برای هر $c \in \mathbb{Q}$ داریم

$$\{w: f(w) + g(w) \leq c\} = \bigcup_{d \in \mathbb{Q}} (\{w: f(w) \leq d\} \cap \{w: g(w) \leq c-d\}) \in F$$

به دلیل اینکه f و g هر دو متغیرند، بنابراین اشتراک دو مجموعه سمت چپ نیز درون F قرار می گیرند.

۹. فرض کنید $E|X| < \infty$ و ثابت کنید $\lim_{x \rightarrow +\infty} xP(X > x) = 0$

پاسخ

$$xP(X > x) = \int_x^\infty x f_x(t) dt \leq \int_x^\infty |x| f_x(t) dt \leq \int_x^\infty |t| f_x(t) dt \leq E(|X|)$$

حال x را به بی نهایت میل می دهیم، به دلیل اینکه آخرین انتگرال کراندار می باشد حد وجود دارد، داریم:



$$0 \leq \lim_{x \rightarrow \infty} xP(X > x) \leq \lim_{x \rightarrow \infty} \int_x^\infty |t| f_x(t) dt = 0$$

$$\lim_{x \rightarrow \infty} xP(X > x) = 0$$

با توجه به قضیه فشردگی نتیجه می‌شود که

$$10. \text{ فرض کنید } E|X|^m < \infty \text{ و ثابت کنید به ازای } 1 \leq m \leq n \text{ نیز } E|X|^m < \infty.$$

پاسخ

برای هر $p, q > 1$ زمانی که در شرط $\frac{1}{p} + \frac{1}{q} = 1$ صدق می‌کنند نامساوی کوشی-شوارتز به فرم کلی زیر تعریف می‌شود:

$$E|XY| \leq \left(E|X|^p\right)^{\frac{1}{p}} \left(E|Y|^q\right)^{\frac{1}{q}}$$

حال اگر در نامساوی کوشی قرار دهیم $X = z^m, Y = 1$, $p = \frac{n}{n-m}, q = \frac{n}{m}$ داریم

$$E(|z|^m) \leq E(|z|^n)^{\frac{m}{n}}$$

به دلیل اینکه $E|z|^n$ متناهی است، متعاقبا $E|z|^m$ نیز متناهی می‌شود.

$$11. \text{ فرض کنید } X_1, X_2, \dots, X_n \text{ متغیرهای تصادفی مستقل و هم توزیع باشند و } S_n = \sum_{i=1}^n X_i \text{ مطلوب است}$$

$$E(S_n | S_m)$$

برای حالت‌های $n < m, n = m, n > m$ بحث کنید.

پاسخ

فرض می‌کنیم $E(X_1) = \mu$ ، آنگاه برای حالت $n = m$

$$E(S_n | S_m) = E(S_n | S_n) = S_n = \sum_{i=1}^n X_i$$

در حالت $n > m$ می‌نویسیم $S_n = S_m + \sum_{i=m+1}^n X_i$ که جمله S_m مستقل از $\sum_{i=m+1}^n X_i$ نیز می‌باشد.

بنابراین:

$$E(S_n | S_m) = E\left(S_m + \sum_{i=m+1}^n X_i | S_m\right) = E(S_m | S_m) + E\left(\sum_{i=m+1}^n X_i | S_m\right) \\ = S_m + (n-m)\mu$$

اگر $n > m$

$$f_{S_n}(s) f_{\sum_{i=m+1}^n X_i}(t-s) \\ f_{S_n | S_m=t}(s) = \frac{f_{S_m}(t)}{f_{S_m}(t)}$$

بنابراین:

$$E(S_n | S_m = t) = \int s \cdot f_{S_n | S_m=t}(s) ds = t \frac{n}{m}$$

راه حل دوم) با استفاده از هم توزیعی متغیرها می‌توان نوشت:

$$E(X_1 | S_m) = E(X_2 | S_m) = \dots = E(X_m | S_m)$$

$$t = E(S_m | S_m = t) = E\left(\sum_{i=1}^m X_i | S_m = t\right) = \sum_{i=1}^m E(X_i | S_m = t) = \sum_{i=1}^m E(X_1 | S_m = t) \\ = mE(X_1 | S_m = t)$$

که از این مساوی آخر نتیجه می‌شود که

$$E(X_1 | S_m = t) = \frac{t}{m}$$

حال اگر $n < m$ باشد، آنگاه

$$E(S_n | S_m) = \sum_{i=1}^n E(X_i | S_m = t) = \sum_{i=1}^n E(X_1 | S_m = t) = n \frac{t}{m}$$

۱۲. فرض کنید $X_1, X_2, \dots$ دنباله‌ای مستقل و هم توزیع با توزیع یکنواخت بر $(0, 1)$ باشند و $E_n = \{n^\alpha X_n > 1\}$.

۱۳. مطلوب است احتمال آنکه E_n ها بی‌نهایت بار رخ دهند، روی مقادیر مختلف α بحث کنید.

پاسخ

E_n ها از یکدیگر مستقلند، زیرا X_n ها مستقلند. برای پاسخگویی به سوال باید احتمال E_n ها را به دست آورد.

$$P(E_n) = P(n^\alpha X_n > 1) = P\left(X_n > \frac{1}{n^\alpha}\right) = 1 - \frac{1}{n^\alpha}$$

اگر $\alpha \leq 0$ باشد، آنگاه برای هر n ، $P(E_n) = 0$ ، یعنی:

$$P(E_n \text{ i.o.}) = 0$$

در حالتی که $\alpha > 1$ در نظر گرفته شده باشد $\sum_{n=1}^{\infty} \left(1 - \frac{1}{n^\alpha}\right) = \infty$ می‌شود، بنابراین بنا به لم بورل-کنتالی

$$P(E_n \text{ i.o.}) = 0$$

در حالتی که $0 < \alpha \leq 1$ باشد، نامساوی $\sum_{n=1}^{\infty} \frac{1}{n^\alpha} < \sum_{n=1}^{\infty} \left(1 - \frac{1}{n^\alpha}\right)$ برقرار می‌باشد. یعنی $\sum_{n=1}^{\infty} \left(1 - \frac{1}{n^\alpha}\right)$ در این حالت

نیز بی‌نهایت می‌شود. بنابراین با بهره‌گیری از لم بورل کنتالی دوباره داریم:

$$P(E_n \text{ i.o.}) = 1$$

و در حالت کلی می‌توان نوشت

$$P(E_n \text{ i.o.}) = 1 \quad \alpha > 0$$

$$P(E_n \text{ i.o.}) = 0 \quad \alpha \leq 0$$



استنباط آماری

۱۴. فرض کنید $X_1, \dots, X_n$ نمونه n تایی از توزیع یکنواخت بر $(0, 1)$ باشد و $X_{(1)} = \min\{X_1, \dots, X_n\}$ و $X_{(n)} = \max\{X_1, \dots, X_n\}$ مطلوب است محاسبه $E(X_{(1)} | X_{(n)})$.

پاسخ

راه حل اول) می‌دانیم که هرگاه $X_i \sim U(a, b)$ آنگاه $\frac{X_i - a}{b - a} \sim U(0, 1)$ است. از طرفی هرگاه $X_i \sim U(0, \theta)$ آنگاه

$X_i \sim U(0, 1)$ و $\frac{X_{(1)}}{\theta} \sim \text{Beta}(1, n)$ و $\frac{X_{(n)}}{\theta} \sim \text{Beta}(n, 1)$ برای θ یک آمار بسنده کامل است. در این تمرین

بنابراین:

$$X_{(n)} \sim \text{Beta}(n, 1) \Rightarrow E(X_{(n)}) = \frac{n}{n+1}$$

$$X_{(1)} \sim \text{Beta}(1, n) \Rightarrow E(X_{(1)}) = \frac{1}{n+1}$$

در نتیجه $\frac{X_{(1)}}{X_{(n)}}$ دارای توزیعی است که به θ بستگی ندارد، یعنی یک آماره کمکی است. از طرفی $X_{(n)}$ آمار بسنده کامل است.

پس با استفاده از قضیه باسو، $\frac{X_{(1)}}{X_{(n)}}$ و $X_{(n)}$ از هم مستقلند. بنابراین:

$$E(X_{(1)} | X_{(n)}) = E\left(\frac{X_{(1)}}{X_{(n)}} \times X_{(n)} | X_{(n)}\right) = E\left(\frac{X_{(1)}}{X_{(n)}} | X_{(n)}\right) \cdot E(X_{(n)} | X_{(n)})$$

$$E\left(\frac{X_{(1)}}{X_{(n)}}\right) \cdot X_{(n)} = \frac{1}{n} \cdot X_{(n)} = \frac{X_{(n)}}{n}$$

نحوه محاسبه $E\left(\frac{X_{(1)}}{X_{(n)}}\right)$ بدین منظور به صورت زیر عمل می‌کنیم.

$$E(X_{(1)}) = E\left(\frac{X_{(1)}}{X_{(n)}} X_{(n)}\right) = E\left(\frac{X_{(1)}}{X_{(n)}}\right) \cdot E(X_{(n)})$$

در نتیجه:

$$E\left(\frac{X_{(1)}}{X_{(n)}}\right) = \frac{E(X_{(1)})}{E(X_{(n)})} = \frac{\frac{1}{n+1}}{\frac{n}{n+1}} = \frac{1}{n}$$

راه حل دوم) با توجه به اطلاعات قبلی می‌دانیم که تابع چگالی آماره ترتیبی و چگالی توام آماره‌های ترتیبی به صورت زیر می‌باشند.

$$F_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} (F(x))^{i-1} (1-F(x))^{n-i} f(x)$$

$$F_{X_{(i)}, X_{(j)}}(x, y) = \frac{n!}{(i-1)!(j-i-1)(n-j)!} (F(x))^{i-1} (F(y)-F(x))^{j-i-1} (1-F(y))^{n-j} f(x) f(y)$$

در این مثال X_i ها دارای توزیع یکنواخت صفر و یک می‌باشند، بنابراین:

$$f(x) = 1 \quad 0 < x < 1$$

$$F(x) = x \quad 0 < x < 1$$

با توجه به مطالب بیان شده تابع چگالی $X_{(n)}$ و توام $X_{(i)}$ و $X_{(j)}$ به صورت زیر می‌باشد.

$$F_{X_{(n)}}(x) = nx^{n-1} \quad 0 < x < 1$$

$$F_{X_{(i)}, X_{(n)}}(x, y) = n(n-1)(y-x)^{n-2} \quad 0 < x < y < 1$$

اگر از تعریف تابع چگالی شرطی استفاده کنیم برای $0 < x < y < 1$ می‌توان نوشت

$$f_{X_{(i)}|X_{(n)}}(x|y) = \frac{f_{X_{(i)}, X_{(n)}}(x, y)}{f_{X_{(n)}}(y)} = \frac{n(n-1)(y-x)^{n-2}}{ny^{n-1}} = \frac{n-1}{y} \left(1 - \frac{x}{y}\right)^{n-2}$$

حال امید ریاضی شرطی متغیر تصادفی $X_{(i)}$ به صورت زیر محاسبه می‌شود.

$$E(X_{(i)}|X_{(n)}=y) = \int_0^y x \frac{n-1}{y} \left(1 - \frac{x}{y}\right)^{n-2} dx$$

با استفاده از روش انتگرال جزء به جزء داریم

$$= -x \left(1 - \frac{x}{y}\right)^{n-1} \Big|_0^y + \int_0^y \left(1 - \frac{x}{y}\right)^{n-1} dx = -x \left(1 - \frac{x}{y}\right)^{n-1} \Big|_0^y - \frac{y}{n} \left(1 - \frac{x}{y}\right)^n \Big|_0^y = \frac{y}{n}$$

یعنی:

$$E(X_{(i)}|X_{(n)}) = \frac{X_{(n)}}{n}$$

۱۵. فرض کنید ظرفی شامل a مهره سفید، b مهره سیاه و c مهره قرمز باشد، از این ظرف n مهره با جای گذاری خارج

می‌کنیم، اگر X تعداد مهره‌های سفید و Y تعداد مهره‌های سیاه خارج شده باشد ضریب همبستگی X, Y را حساب کنید؟

پاسخ

به دلیل اینکه روش با جای گذاری انتخاب شده، بنابراین:

$$P(\text{انتخاب مهره سفید}) = \frac{a}{a+b+c} = p_1$$

$$P(\text{انتخاب مهره سیاه}) = \frac{b}{a+b+c} = p_2$$

$$P(\text{انتخاب مهره قرمز}) = \frac{c}{a+b+c} = p_3$$

انتخاب n مهره با جای گذاری، یک سه جمله‌ای می‌سازد. بنابراین:

$$\text{Cov}(X, Y) = -p_1 p_2 = -\frac{ab}{(a+b+c)^2}$$

به دلیل اینکه توابع چگالی حاشیه‌ای X و Y توزیع دو جمله‌ای می‌باشند، می‌توان نوشت

$$\text{Var}(X) = p_1(1-p_1)$$

$$\text{Var}(Y) = p_2(1-p_2)$$

حال با توجه به این مطالب و مستقل بودن نمونه‌گیری‌های مختلف (استفاده از روش با جای‌گذاری)، ضریب همبستگی X و Y برابر است با:

$$\begin{aligned} \text{Corr}(X, Y) &= \rho_{x,y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{-p_1 p_2}{\sqrt{p_1 p_2 (1-p_1)(1-p_2)}} \\ &= -\sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}} \end{aligned}$$

۱۶. خانواده توزیع‌های کامل را تعریف کنید. یک خانواده کامل و یک خانواده که کامل نباشد مثال بزنید. ادعاهای خود را ثابت کنید.

پاسخ

خانواده توزیع‌های تولیدشده توسط آماره T ، $\{p_\theta^T, \theta \in \Theta\}$ را کامل گوئیم اگر برای هر آماره $g(T)$ که امید ریاضی آن برای هر θ عضو فضای پارامتری صفر می‌باشد نتیجه بگیریم که یقیناً این آمار مقدار صفر می‌پذیرد، یعنی

$$E(g(T)) = 0 \quad \forall \theta \in \Theta \Rightarrow p(g(T) = 0) = 1 \quad \forall \theta \in \Theta$$

مثال برای کامل بودن

فرض کنید نمونه $X_1, \dots, X_n$ از توزیع پواسون با پارامتر θ باشد. نشان دهید که خانواده توزیع تولیدشده توسط $T = \sum_{i=1}^n x_i$ کامل است.

$$X_1, \dots, X_n \sim \text{pos}(\theta) \quad \theta > 0$$

$$T = \sum_{i=1}^n X_i \sim \text{pos}(n\theta)$$

از طرفی $T = \sum_{i=1}^n X_i$ برای θ یک آمار بسنده است. حال اگر برای هر θ ، $E(g(T)) = 0$ باشد، آنگاه

$$E(g(T)) = \sum_{t=0}^{\infty} g(t) \frac{e^{-n\theta} (n\theta)^t}{t!} = 0$$

فرض می‌کنیم $h(t) = \frac{n^t}{t!} g(t)$ ، آنگاه

$$\sum_{t=0}^{\infty} \underbrace{g(t) \frac{n^t}{t!}}_{h(t)} \theta^t = \sum_{t=0}^{\infty} h(t) \theta^t = 0$$

که این یک سری توانی از θ است و زمانی صفر می‌شود که برای هر t مقدار $h(t)$ صفر باشد، یعنی برای هر θ

$$g(t) = 0 \quad \forall t \Rightarrow p(g(T) = 0) = 1 \quad \forall \theta \in \Theta$$

مثال برای کامل بودن

فرض کنید نمونه $X_1, \dots, X_n$ از توزیع پواسون با پارامتر θ باشد. نشان دهید که خانواده توزیع تولیدشده توسط $T = \sum_{i=1}^n x_i$ کامل است.