آمار و احتمال مهندسی

نیم‌سال دوم ۱۴۰۳-۱۴۰۲

مدرس: دکتر امیر نجفی

دانشکده‌ی مهندسی کامپیوتر

تمرین سری دوم | زمان تحویل: ۲۸ اسفند

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص‌شده است.
- هم‌کاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت هم‌فکری و یا استفاده از هر منابع خارج درسی، نام هم‌فکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

بارم‌بندی

بارم سوالات به شکل زیر است: (مجموعا ۱۰۰ نمره)

- سوال ۱: ۱۲ امتیاز
- سوال ۲: ۱۰ امتیاز
- سوال ۳: ۱۵ امتیاز
- سوال ۴: ۱۲ امتیاز
- سوال ۵: ۱۰ امتیاز
- سوال ۶: ۱۰ امتیاز
- سوال ۷: ۹ امتیاز
- سوال ۸: ۱۰ امتیاز
- سوال ۹: ۱۲ امتیاز

مسئله‌ی ۱. (چگالی و تجمع)

متغیر تصادفی X با تابع چگالی احتمال زیر داده شده است:

$$f_X(x) = \begin{cases} ae^{-\lambda(x-1)}, & x > 1 \\ b, & -1 \leqslant x \leqslant 1 \\ ce^{\lambda(x+1)}, & x < -1 \end{cases}$$

به طوری که داریم: $\lambda > 0$.

الف) ثابت کنید $a + c = \lambda(1 - 2b)$

ب) تابع توزیع تجمعی را برای متغیر تصادفی X محاسبه نمایید.

جواب.

الف)

می دانیم که باید $\int_{-\infty}^{\infty} f_X(x)dx = 1$. همچنین داریم :

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_{-\infty}^{-1} ce^{\lambda(x+1)}\, dx + \int_{-1}^{1} b\, dx + \int_{1}^{\infty} ae^{-\lambda(x-1)}\, dx =$$

$$\frac{c}{\lambda}e^{\lambda(x+1)}\bigg|_{-\infty}^{-1} + bx\bigg|_{-1}^{1} + \frac{-a}{\lambda}e^{-\lambda(x-1)}\bigg|_{1}^{\infty} = \frac{c}{\lambda} + 2b + \frac{-a}{\lambda}(-1) = \frac{a+c}{\lambda} + 2b$$

$$\Rightarrow \frac{a+c}{\lambda} + 2b = 1 \Rightarrow a + c = \lambda(1 - 2b)$$

ب)

میخواهیم $F_X(x)$ را محاسبه کنیم .

میدانیم که $F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$ است. پس روی x حالت بندی می‌کنیم :

حالت ۱ : $x < -1$

$$F_X(x) = \int_{-\infty}^{x} ce^{\lambda(t+1)}\, dt = \frac{c}{\lambda}e^{\lambda(t+1)}\bigg|_{-\infty}^{x} = \frac{c}{\lambda}e^{\lambda(x+1)}$$

حالت ۲ : $-1 \leqslant x \leqslant 1$

$$F_X(x) = \int_{-\infty}^{-1} ce^{\lambda(t+1)}\, dt + \int_{-1}^{x} b\, dt = \frac{c}{\lambda}e^{\lambda(t+1)}\bigg|_{-\infty}^{-1} + bt\bigg|_{-1}^{x} = \frac{c}{\lambda} + b(x+1)$$

حالت ۳ : $x > 1$

$$F_X(x) = \int_{-\infty}^{-1} ce^{\lambda(t+1)}\, dt + \int_{-1}^{1} b\, dt + \int_{1}^{x} ae^{-\lambda(t-1)}\, dt =$$

$$\frac{c}{\lambda}e^{\lambda(t+1)}\bigg|_{-\infty}^{-1} + bt\bigg|_{-1}^{1} + \frac{-a}{\lambda}e^{-\lambda(t-1)}\bigg|_{1}^{x} = \frac{c}{\lambda} + 2b - \frac{a}{\lambda}(e^{-\lambda(x-1)} - 1) =$$

$$\frac{a+c}{\lambda} + 2b - \frac{a}{\lambda}e^{-\lambda(x-1)} = 1 - \frac{a}{\lambda}e^{-\lambda(x-1)}$$

$$\Rightarrow F_X(x) = \begin{cases} \frac{c}{\lambda}e^{\lambda(x+1)}, & x < -1 \\ \frac{c}{\lambda} + b(x+1), & -1 \leqslant x \leqslant 1 \\ 1 - \frac{a}{\lambda}e^{-\lambda(x-1)}, & 1 < x \end{cases}$$

مسئله‌ی ۲. (نخ‌های گره خورده)

فرض کنید n نخ داریم. می‌دانیم که هر نخ دو سر دارد؛ لذا کل نخ‌ها $2n$ سر خواهند داشت. به صورت تصادفی n زوج از سرهای این نخ‌ها را در نظر می‌گیریم و دو سرِ تشکیل‌دهنده‌ی هر زوج را به هم گره می‌زنیم. اگر تعداد حلقه‌های ایجادشده را با یک متغیر تصادفی مانند L نشان دهیم، مطلوب است $\mathbb{E}[L]$.

جواب.

پاسخ مسئله برای n نخ را با e_n نمایش می دهیم. به وضوح $e_1 = 1$ می باشد. فرض کنید مسئله را برای $n-1$ نخ حل کرده ایم و حال یک نخ دیگر اضافه می کنیم. یک سر این نخ را در نظر بگیرید؛ دو حالت برای آن امکان پذیر است:

۱- به سر دیگر همین نخ متصل شود. احتمال رخ دادن این رویداد، $\frac{1}{2n-1}$ می باشد. در این حالت میانگین تعداد حلقه های تولید شده توسط سایر نخ ها، مستقل از این نخ، برابر e_{n-1} می باشد. پس میانگین تعداد کل حلقه های تولید شده در این حالت برابر $e_{n-1} + 1$ خواهد بود.

۲- به یک سر از یک نخ دیگر متصل شود. احتمال رخ دادن این رویداد، $\frac{2n-2}{2n-1}$ می باشد. در این حالت می توانیم سر های آزاد این دو نخ را به عنوان دو سر یک نخ جدید در نظر بگیریم و خود این دو نخ را حذف کنیم. پس در ادامه $n-1$ نخ خواهیم داشت که از اتصال دو به دوی سر های آن ها به طور میانگین e_{n-1} حلقه تولید خواهد شد.

در نتیجه e_n از رابطه زیر به صورت بازگشتی محاسبه می شود:

$$e_n = \frac{1}{2n-1}(e_{n-1}+1) + \frac{2n-2}{2n-1}e_{n-1} = \frac{1}{2n-1} + e_{n-1}$$

$$\Longrightarrow e_n = 1 + \frac{1}{3} + \frac{1}{5} + ... + \frac{1}{2n-1}$$

مسئله‌ی ۳. (میانگین و واریانس)

- T یک متغیر تصادفی گسسته است. به طوری که تابع جرم احتمال آن به صورت زیر تعریف می‌شود:

$$\forall n \in \mathbb{N} \quad \mathbb{P}(T = n) = \frac{A}{n \cdot ۳^n},$$

که در آن A یک مقدار ثابت است.
الف) مقدار عددی ثابت A را بیابید.
ب) مقادیر $\mathbb{E}[T]$ و $\mathrm{Var}[T]$ را محاسبه کنید.

- متغیر تصادفی X را در نظر بگیرید به طوری که

$$\forall n \in N \quad \mathbb{P}(X = n) = \frac{c}{n!}$$

مقدار c را بیابید و سپس مقادیر $\mathbb{E}[X]$ و $\mathrm{Var}[X]$ را به دست آورید.

جواب.

- بخش اول:

الف)

$$\forall ۰ < p < ۱ \; ; \; \sum_{n=۰}^{\infty} p^n = \frac{۱}{۱ - p} \implies \int_۰^p \sum_{n=۰}^{\infty} x^n \, dx = \int_۰^p \frac{۱}{۱ - x} \, dx$$

$$\implies \sum_{n=۱}^{\infty} \frac{p^n}{n} = \ln(\frac{۱}{۱ - p}) \implies \sum_{n=۱}^{\infty} \frac{۱}{n \cdot ۳^n} = \ln(\frac{۳}{۲})$$

در نتیجه برای مقدار A داریم:

$$\sum_{n=۰}^{\infty} P(T = n) = ۱ \implies \sum_{n=۰}^{\infty} \frac{A}{n \cdot ۳^n} = A \ln(\frac{۳}{۲}) = ۱ \implies A = \frac{۱}{\ln(\frac{۳}{۲})}$$

ب)

$$\mathbb{E}[T] = \sum_{n=۱}^{\infty} n P(T = n) = \frac{۱}{\ln(\frac{۳}{۲})} \sum_{n=۱}^{\infty} \frac{۱}{۳^n} = \frac{۱}{۲ \ln(\frac{۳}{۲})}$$

حال دقت کنید که می توان نوشت:

$$\forall ۰ < p < ۱ \; ; \; \sum_{n=۰}^{\infty} p^n = \frac{۱}{۱ - p} \implies \frac{d}{dp} \sum_{n=۰}^{\infty} p^n = \frac{d}{dp} \frac{۱}{۱ - p}$$

$$\implies \sum_{n=۱}^{\infty} n p^{n-۱} = \frac{۱}{(۱ - p)^۲} \implies \sum_{n=۱}^{\infty} n p^n = \frac{p}{(۱ - p)^۲} \implies \sum_{n=۱}^{\infty} \frac{n}{۳^n} = \frac{۳}{۴}$$

پس برای محاسبه واریانس داریم:

$$\mathbb{E}[T^۲] = \sum_{n=۱}^{\infty} n^۲ P(T = n) = \frac{۱}{\ln(\frac{۳}{۲})} \sum_{n=۱}^{\infty} \frac{n}{۳^n} = \frac{۳}{۴\ln(\frac{۳}{۲})}$$

$$Var[T] = \mathbb{E}[T^۲] - \mathbb{E}[T]^۲ = \frac{۳}{۴\ln(\frac{۳}{۲})} - \frac{۱}{۴\ln(\frac{۳}{۲})^۲}$$

- بخش دوم:

$$\sum_{n=۱}^{\infty} P(X = n) = ۱ \Longrightarrow \sum_{n=۱}^{\infty} \frac{c}{n!} = c(e - ۱) = ۱ \Longrightarrow c = \frac{۱}{e - ۱}$$

$$\mathbb{E}[X] = \sum_{n=۱}^{\infty} nP(X = n) = \frac{۱}{e - ۱} \sum_{n=۱}^{\infty} \frac{۱}{(n - ۱)!} = \frac{e}{e - ۱}$$

$$\mathbb{E}[X^۲] = \sum_{n=۱}^{\infty} n^۲ P(X = n) = \frac{۱}{e - ۱} \sum_{n=۱}^{\infty} \frac{n}{(n - ۱)!} = \frac{۱}{e - ۱} \left(۱ + \sum_{n=۲}^{\infty} \frac{۱}{(n - ۱)!} + \frac{۱}{(n - ۲)!} \right) = \frac{۲e}{e - ۱}$$

$$Var[X] = \mathbb{E}[X^۲] - \mathbb{E}[X]^۲ = \frac{۲e}{e - ۱} - \left(\frac{e}{e - ۱} \right)^۲ = \frac{e^۲ - ۲e}{(e - ۱)^۲}$$

مسئله‌ی ۴. (قدِ نرمال)

فرض کنید قد مردان ۲۵ ساله یک متغیر تصادفی است که از توزیع نرمال با میانگین ۱۷۰ و واریانس ۱۰ پیروی کند.

الف) قد چند درصد از این مردان بیش از ۱۷۴ سانتی‌متر است؟

ب) چند درصد از مردانی که قد بیشتر یا مساوی ۱۶۰ سانتی‌متر دارند، قدشان کمتر از ۱۸۰ سانتی‌متر است؟

پ) اگر بدانیم که قد افراد مختلف مستقل است و از همین توزیع داده‌شده در سوال پیروی می‌کند (به بیان دیگر، اعداد متناظر قد آن‌ها متغیرهای تصادفی i.i.d باشند)، چقدر احتمال دارد که از میان یک گروه ۵ نفره از این مردان، حداقل نصف آن‌ها بیش از ۱۷۴ سانتی‌متر قد داشته باشند؟

جواب.

می دانیم که قد مردان ۲۵ ساله یک متغیر تصادفی با میانگین ۱۷۰ و واریانس ۱۰ است. بنابراین برای توزیع نرمال داریم:

$$f_X(x = a) = \frac{۱}{\sigma\sqrt{۲\pi}} e^{\frac{-۱}{۲}(\frac{a-\mu}{\sigma})^۲}$$

الف) به متغیر تصادفی نرمال استاندارد زیر تغییر متغیر می‌دهیم:

$$Z = \frac{X-\mu}{\sigma} = \frac{X-۱۷۰}{\sqrt{۱۰}}$$

$$P(X > ۱۷۴) = ۱ - P(X \leqslant ۱۷۴) \Longrightarrow P(X \leqslant ۱۷۴) = P(X - ۱۷۰ \leqslant ۱۷۴ - ۱۷۰)$$
$$= P(\frac{X-۱۷۰}{\sqrt{۱۰}} \leqslant \frac{۱۷۴-۱۷۰}{\sqrt{۱۰}}) = P(Z \leqslant \frac{۴}{\sqrt{۱۰}}) = \Phi(\frac{۴}{\sqrt{۱۰}}) = ۰/۸۹$$
$$P(X > ۱۷۴) = ۱ - ۰/۸۹ = ۰/۱۱$$

ب) برای محاسبه ابتدا باید درصد افرادی که قد بالای ١٦٠ دارند را حساب کنیم و سپس افرادی که قد کمتر از ١٨٠ دارند:

$$P(X > ۱۶۰) = ۱ - P(X \leqslant ۱۶۰) \Longrightarrow P(X \leqslant ۱۶۰) = P(X - ۱۷۰ \leqslant ۱۶۰ - ۱۷۰)$$
$$= P(\frac{X-۱۷۰}{\sqrt{۱۰}} \leqslant \frac{۱۶۰-۱۷۰}{\sqrt{۱۰}}) = P(Z \leqslant \frac{-۱۰}{\sqrt{۱۰}}) = \Phi(\frac{-۱۰}{\sqrt{۱۰}}) = ۰/۰۰۰۷$$
$$P(X > ۱۶۰) = ۱ - ۰/۰۰۰۷ = ۰/۹۹۹۳$$
$$P(X < ۱۸۰) = P(X - ۱۷۰ < ۱۸۰ - ۱۷۰) = P(\frac{X-۱۷۰}{\sqrt{۱۰}} < \frac{۱۸۰-۱۷۰}{\sqrt{۱۰}})$$
$$= \Phi(\frac{۱۰}{\sqrt{۱۰}}) = ۰/۹۹۹۲ \Longrightarrow P(X >= ۱۸۰) = ۱ - ۰/۹۹۹۲ = ۰/۰۰۰۸$$

پس نتیجه می گیریم که ۰/۹۹۹۳ افراد دارای قد بالای ١٦٠ سانتی متر هستند و ۰/۰۰۰۸ افراد جامعه دارای قد بالای ١٨٠ سانتی متر هستند بنابراین

$$۰/۹۹۹۳ - ۰/۰۰۰۸ = ۰/۹۹۸۵ \Longrightarrow \frac{۰/۹۹۸۵}{۰/۹۹۹۳} = ۰/۹۹۹۱$$

بنابراین نشان می دهد که ۹۹/۹۱ درصد افرادی که قد بالای ١٦٠ دارند قدشان زیر ١٨٠ است.

پ) ابتدا حالت‌های مختلف که ممکن است بوجود بیاید را بررسی می کنیم و بعد از آن احتمال کل را جمع می کنیم:

حالت اول:سه نفر دارای قد بیشتر از ١٧٤ سانتی متر باشند و ٢ نفر کمتر از ١٧٤ سانتی متر:

$$\binom{۵}{۳} P(X > ۱۷۴)^{۳} P(X < ۱۷۴)^{۲} = ۱۰ \times (۰/۱۱)^{۳} \times (۰/۸۹)^{۲} = ۰/۰۱۰۵$$

حالت دوم:چهار نفر دارای قد بیشتر از ١٧٤ سانتی متر باشند و یکی کمتر:

$$\binom{۵}{۴} P(X > ۱۷۴)^{۴} P(X < ۱۷۴)^{۱} = ۱۰ \times (۰/۱۱)^{۴} \times (۰/۸۹)^{۱} = ۰/۰۰۰۶۵$$

حالت سوم:هر پنج نفر دارای قد بیشتر از ۱۷۴ سانتی متر باشند:

$$\binom{۵}{۵} P(X > ۱۷۴)^{۵} P(X < ۱۷۴)^{۰} = ۱ \times (۰/۱۱)^{۵} \times (۰/۸۹)^{۰} = ۰/۰۰۰۰۶۱$$

پس احتمال آن که در این جمع بیش از نیمی از افراد دارای قد بیشتر از ۱۷۴ سانتی متر باشند برابر است با:

$$۰/۰۱۰۵ + ۰/۰۰۰۶۵ + ۰/۰۰۰۰۶۱ = ۰/۰۱۱۲$$

مسئله‌ی ۵. (پواسون)

اگر X و Y دو متغیر تصادفی پواسون با پارامتر $\lambda = ۲$ باشند و Z مینیمم آن‌ها را نشان دهد، مقدار $\mathbb{P}(Z \leqslant ۱)$ را محاسبه کنید.

جواب.

می دانیم که $Z = min\{X, Y\}$ است. CDF متفیر بنابراین محاسبه مستقیم Z دشوار است و به جای بدست آوردن آن، احتمال متمم آن را بدست می آوریم:

$$P(Z \leqslant ۱) = ۱ - P(Z > ۱)$$

$$P(Z > ۱) = P(min\{X, Y\} > ۱) = P(X > ۱, Y > ۱)$$

می دانیم که متغیرهای تصادفی X و Y مستقل از هم هستند و توزیع یکسان دارند. بنابراین داریم:

$$P(X > ۱, Y > ۱) = P(X > ۱)P(Y > ۱) = P(X > ۱)^{۲} = (۱ - P(x \leqslant ۱))^{۲}$$

یادآوری: برای توزیع پواسون داریم:

$$P(X = a) = \frac{\lambda^a}{a!} e^{-\lambda}$$

پس برای حل این سوال خواهیم داشت:

$$P(X \leqslant ۱) = \frac{۲^{۰}}{۰!} e^{-۲} + \frac{۲^{۱}}{۱!} e^{-۲} = ۳e^{-۲} \Rightarrow \; P(X > ۱) = ۱ - P(X \leqslant ۱) = ۱ - ۳e^{-۲} \Rightarrow$$

$$P(X > ۱)P(Y > ۱) = P(X > ۱)^{۲} = (۱ - ۳e^{-۲})^{۲} \Rightarrow \; P(Z > ۱) = (۱ - ۳e^{-۲})^{۲} \Rightarrow$$

$$P(Z \leqslant ۱) = ۱ - P(Z > ۱) = ۱ - (۱ - ۳e^{-۲})^{۲}$$

مسئله‌ی ۶. (بی‌حافظگی)

متغیر تصادفی X دارای خاصیت بی‌حافظگی[1] است؛ اگر داشته باشیم:

$$\forall m, n \in \mathbb{N} \cup \{۰\} \quad \mathbb{P}[X > n + m | X > n] = \mathbb{P}[X > m]$$

اگر X یک متغیر تصادفی گسسته باشد، نشان دهید بی‌حافظه است؛ اگر و تنها اگر دارای توزیع هندسی باشد.

جواب.

ابتدا فرض کنید $X \sim Geo(p)$:

$$P[X > n] = ۱ - \sum_{k=۱}^{n} P[X = k] = ۱ - \sum_{k=۱}^{n} pq^{k-۱} = ۱ - p\frac{۱ - q^n}{۱ - q} = q^n$$

$$\Rightarrow P[X > n + m | x > n] = \frac{P[X > n + m]}{P[X > n]} = \frac{q^{n+m}}{q^n} = q^m = P[X > m]$$

حال فرض کنید X بی حافظه باشد .

$$n = m = ۰ \Rightarrow P[X > ۰] = P[X > ۰ | X > ۰] = ۱$$

پس X فقط مقادیر طبیعی می پذیرد .
فرض کنید $q = P[X > ۱]$ باشد.

می‌دانیم:

$$P[X > n + m | X > n] = \frac{P[X > n + m]}{P[X > n]}$$

پس برای هر $n, m \in \mathbb{N}$ داریم:

$$P[X > n + m] = P[X > n].P[X > m]$$

پس می‌توان به صورت استقرایی نتیجه گرفت که:

$$\forall n \in \mathbb{N} : P[X > n] = q^n$$

[1] Memoryless Property

$$\Rightarrow \forall n \in \mathbb{N} : P[X = n] = P[X > n - ۱] - P[X > n] = q^{n-۱} - q^n$$

$$= (۱ - q)q^{n-۱} \Rightarrow X \sim Geo(۱ - q)$$

مسئله‌ی ۷. (چمدانِ پروفسور)

پروفسور سراجی با هواپیما از لس‌آنجلس به پاریس سفر می‌کند. برای این منظور، او ابتدا از فرودگاه لس‌آنجلس به نیویورک و سپس از آنجا به لندن می‌رود و در نهایت از لندن به پاریس می‌رسد. می‌دانیم احتمال آن که وسایل یک مسافر در هر یک از فرودگاه‌های لس‌آنجلس، نیویورک و لندن گم شود، p است. وقتی پروفسور به پاریس می‌رسد، متوجه می‌شود که چمدانش گم شده است. چقدر احتمال دارد که چمدانش در هر یک از فرودگاه‌های مذکور گم شده باشد؟

جواب.

می دانیم که احتمال اینکه وسایلش در هریک از شهرهای مذکور گم شده باشند p و مستقل از سایر شهرها است. بنابراین احتمال اینکه در هر یک از شهر ها وسایلش گم شده باشد دارای توزیع هندسی با پارامتر p است.

یادآوری: در توزیع هندسی احتمال اینکه در kامین آزمایش پیشامد مورد نظر اتفاق بیافتد برابر است با:

$$P(X = k) = (۱ - p)^{k-۱}p$$

پس برای هريك از شهرهای مذکور در بین راه خواهیم داشت:

احتمال آنکه در نیویورگ گم شده باشد:

$$k = ۱ \Rightarrow \; P(X = k) = (۱ - p)^{۱-۱}p = p$$

احتمال آنکه در لندن گم شده باشد:

$$k = ۲ \Rightarrow \; P(X = k) = (۱ - p)^{۲-۱}p = (۱ - p)p$$

احتمال آنکه در پاریس گم شده باشد:

$$k = ۳ \Rightarrow \; P(X = k) = (۱ - p)^{۲-۱}p = (۱ - p)^۲p$$

مسئله‌ی ۸. (نظریه بازی‌ها)

استاد درس نظریه بازی‌ها ایده جالبی برای آزمون میانترم به ذهنش رسیده است. او این آزمون را با k سوال برگزار می‌کند؛ اما یک راه کسب نمره تشویقی نیز برای آن در نظر گرفته است. اگر سوالی فقط توسط یک نفر حل شود، نمره آن برای او دو برابر حساب می‌شود.

می‌دانیم احتمال حل‌شدن هر سوال توسط هر فرد p است و این رویداد، هیچ‌گونه وابستگی به بقیه سوالات و افراد ندارد. استاد درس، برای هر سوال ۱ نمره در نظر گرفته و در زمان تصحیح، ممکن است کل نمره یا صفر را به پاسخ هر دانشجو اختصاص دهد (دقت کنید که حالت دیگری برای نمره‌دهی وجود ندارد). اگر کلاس n نفره باشد، امید ریاضی میانگین نمره کلاس در این آزمون را به دست آورید.

جواب.

طبق خاصیت خطی بودن امید ریاضی، می توان هر سوال را به صورت جداگانه بررسی کرد. حال برای سوال دلخواه، متغیر تصادفی X_i را برابر نمره دانشجوی i ام قبل از اعمال نمره امتیازی تعریف می کنیم. همچنین قرار می دهیم:

$$P_i = \begin{cases} ۱ & X_i = ۱ \ \wedge \ \forall j \neq i \ ; \ X_j = ۰ \\ ۰ & otherwise \end{cases}$$

دقت کنید که برای محسابه $\mathbb{E}[P_i]$ می توان نوشت:

$$\mathbb{E}[P_i] = p(۱-p)^{n-۱}$$

چرا که تنها حالت مطلوب در فضای حالات، این است که فقط شخص i ام این سوال را حل کند.
درنتیجه برای میانگین این سوال خواهیم داشت:

$$Y = \frac{\sum_{i=۱}^{n} X_i + P_i}{n}$$

حال داریم:

$$\mathbb{E}[Y] = \mathbb{E}\left[\ \frac{\sum_{i=۱}^{n} X_i + P_i}{n}\right] \ = \mathbb{E}[X_۱] + \mathbb{E}\left[\ \frac{\sum_{i=۱}^{n} P_i}{n}\right] \ = \mathbb{E}[X_۱] + \mathbb{E}[P_۱] = p + p(۱-p)^{n-۱}$$

پس میانگین کل آزمون برابر خواهد بود با:

$$kp(۱ + (۱-p)^{n-۱})$$

مسئله‌ی ۹. (چندجمله‌ای)

تابع توزیع تجمعی (CDF) متغیر تصادفی پیوسته X به صورت زیر داده شده است:

$$F_X(x) = \begin{cases} ۰ & x < -۱ \\ Q(x) & -۱ \leqslant x \leqslant ۲ \\ ۱ & x > ۲ \end{cases}$$

که در آن $Q(x)$ یک چند جمله ای از درجه ۳ است. اگر بدانیم $P(۰ \leqslant X) = \frac{۲۶}{۲۷}$ و $P(X \leqslant ۱) = \frac{۸}{۲۷}$ ، آنگاه:
الف) $Q(x)$ را مشخص نمایید.
ب) تابع چگالی احتمال متغیر تصادفی X را به دست آورید.

جواب.

الف) می دانیم که پیوستگی یکی از خواص تابع توزیع تجمعی (CDF) می باشد. با توجه به شرط پیوستگی برای تابع F_x در نقاط $x = -۱$ و $x = ۲$ داریم:

$$Q(-۱) = ۰\ ,\ Q(۲) = ۱$$

همچنین با استفاده از داده های سوال، می توان نوشت:

$$P(۰ \leqslant X) = ۱ - P(X < ۰) = ۱ - F_X(۰) = \frac{۲۶}{۲۷} \Longrightarrow F_X(۰) = Q(۰) = \frac{۱}{۲۷}$$

$$P(X \leqslant ۱) = F_X(۱) = Q(۱) = \frac{۸}{۲۷}$$

پس برای چند جمله ای $Q(x)$ داریم:

$$Q(x) = \frac{(x+۱)(x-۱)(x-۲)}{۲} \times \frac{۱}{۲۷} + \frac{x(x+۱)(x-۲)}{-۲} \times \frac{۸}{۲۷} + \frac{x(x+۱)(x-۱)}{۶} = \frac{(x+۱)^{۳}}{۲۷}$$

ب) با مشتق گیری از تابع توزیع تجمعی، تابع چگالی احتمال متغیر تصادفی X را می یابیم:

$$f_X(x) = \begin{cases} ۰ & x < -۱ \\ \frac{(x+۱)^{۲}}{۹} & -۱ \leqslant x \leqslant ۲ \\ ۰ & x > ۲ \end{cases}$$