حل سوالات میان‌ترم اول

آمار و احتمالات مهندسی
اردیبهشت ۱۴۰۰

نگارش: امیر نجفی

توضیحات: جواب‌ها در این پاسخ‌نامه تعمداً با تفصیل و توضیحات زیاد ارائه شده‌اند. اما برای اخذ نمره سوال هنگام تصحیح صرف انتقال مفهوم کافی است.

سوال ۱ (۳ نمره):

یک فرستنده قصد ارسال یک بیت اطلاعات را به یک گیرنده دارد. این بیت را با $X \in \{0,1\}$ نشان می‌دهیم. در هنگام دریافت و به دلیل وجود نویز، گیرنده به جای دریافت X، مقدار تخریب شده $X+Z$ را دریافت می‌کند. در اینجا $Z \sim \mathrm{Exp}(\lambda)$ یک متغیر تصادفی نمایی مستقل از X است. گیرنده که یک عدد پیوسته را دریافت کرده، به طریقه زیر از عدد دریافتی رمزگشایی می‌کند: در صورتیکه مقدار دریافت شده از ۱ کوچک‌تر بود بیت دریافتی را ۰ تلقی کرده، و در غیر این صورت آن را ۱ محسوب می‌کند.

الف) در یک یا دو خط استدلال کنید که در صورتیکه بیت ارسالی $X=1$ باشد، در گیرنده بدون خطا تشخیص داده می‌شود.

ب) به شرط اینکه بدانیم بیت ارسالی مقدار واقعی $X=0$ را داشته، احتمال بروز خطا در گیرنده چقدر است؟

ج) فرض کنید توزیع X یک برنولی با احتمال موفقیت p است: یعنی $X \sim \mathrm{Bern}\left(p\right)$. احتمال اینکه گیرنده بیت ارسالی را به اشتباه تشخیص دهد چقدر است؟

جواب:

الف) توزیع نمایی دارای چگالی $\lambda e^{-\lambda x}u(x)$ بوده و احتمال اینکه مقادیر منفی به خود بگیرد صفر است. لذا $X+Z$ همواره از X بزرگ‌تر است. اگر بدانیم $X=1$ بوده است، عدد دریافت شده در سمت گیرنده نیز بزرگ‌تر از ۱ خواهد بود و توسط رمزگشا به ۱ نگاشت داده خواهد شد. لذا خطایی رخ نمی‌دهد.

ب) در صورتیکه بدانیم $X=0$ بوده است، در گیرنده $Z \sim \lambda e^{-\lambda z}u(z)$ را دریافت کرده‌ایم. از طرفی خطا در این حالت زمانی رخ خواهد داد که رمزگشا Z را به ۱ نگاشت دهد، یعنی داشته باشیم $Z \geq 1$. در نتیجه خواهیم داشت:

$$\mathbb{P}\left(\text{error}\,|\,X=0\right) = \mathbb{P}(Z \geq 1) = \int_1^\infty \lambda e^{-\lambda z}\mathrm{d}z = e^{-\lambda}$$

پس جواب $e^{-\lambda}$ است.

ج) طبق قانون احتمال کل داریم:

$$\mathbb{P}\,(\text{error}) = \mathbb{P}\left(\text{error}\,|\,X=0\right)(1-p) + \mathbb{P}\left(\text{error}\,|\,X=1\right)p = (1-p)e^{-\lambda}$$

لذا جواب $(1-p)e^{-\lambda}$ است.

سوال ۲ (۴ نمره):

الف) فرض کنید متغیر تصادفی X توزیع یکنواخت بین ۰ و ۱، و متغیر تصادفی Y توزیع $\mathrm{Exp}(\lambda)$ داشته باشند. همچنین این دو از هم مستقل هستند. مقدار $\mathbb{P}\,(Y \geq X)$ را محاسبه کنید.

ب) فرض کنید $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ و $Y \sim \mathcal{N}\left(\mu', \sigma'^2\right)$ دو متغیر تصادفی مشترکاً گاوسی با ضریب همبستگی r باشند. همچنین، می‌دانیم که برای دو متغیر تصادفی مشترکاً گاوسی با مشخصات فوق حتماً داریم: $\mathbb{E}\left[(X-\mu)(Y-\mu')\right] = r\sigma\sigma'$
در این صورت، توزیع دقیق $Z = 2X + 3Y$ را بیابید.

جواب:

الف) با توجه به مفروضات سوال، و قانون احتمال کل و تعاریف احتمالات شرطی داریم:

$$\mathbb{P}(Y > X) = \int_{-\infty}^{+\infty} \mathbb{P}(Y > x) f_X(x)\mathrm{d}x = \int_0^1 \mathbb{P}(Y > x)\mathrm{d}x$$

از طرفی می‌دانیم $\mathbb{P}(Y > x) = 1 - F_Y(x) = e^{-\lambda x}$ است. لذا در نهایت خواهیم داشت:

$$\mathbb{P}(Y > X) = \int_0^1 e^{-\lambda x}\mathrm{d}x = \frac{1-e^{-\lambda}}{\lambda}.$$

پس جواب نهایی $(1-e^{-\lambda})/\lambda$ است.

ب) چون X, Y مشترکاً گاوسی هستند، هر ترکیب خطی از آنان توزیع گاوسی خواهد داشت. در نتیجه داریم: $Z \sim \mathcal{N}(A, B)$. پس می‌ماند تعیین مقدار میانگین A و واریانس B تا توزیع دقیقاً مشخص شود. برای مقدار میانگین داریم:

$$A = \mathbb{E}(Z) = \mathbb{E}(2X + 3Y) = 2\mathbb{E}(X) + 3\mathbb{E}(Y) = 2\mu + 3\mu'$$

از طرفی برای $B = \mathrm{Var}(Z)$ خواهیم داشت:

$$B = \mathbb{E}\left[(2X + 3Y - 2\mu - 3\mu')^2\right] = 2^2\mathbb{E}\left[(X-\mu)^2\right] + 3^2\mathbb{E}\left[(Y-\mu')^2\right] + 2\cdot 2\cdot 3\mathbb{E}\left[(X-\mu)(Y-\mu')\right]$$

که طبق تعریف واریانس، و همچنین گفته سوال درباره $\mathbb{E}\left[(X-\mu)(Y-\mu')\right] = r\sigma\sigma'$ برابر خواهد بود با:

$$B = 4\sigma^2 + 9\sigma'^2 + 12r\sigma\sigma'$$

و در نهایت خواهیم داشت:

$Z \sim \mathcal{N}(2\mu + 3\mu' \ , 4\sigma^2 + 9\sigma'^2 + 12r\sigma\sigma')$.

سوال ۳ (۴ نمره):

یک شرکت بیمه n مشترک دارد (فرض کنید $n \gg 1$). قرارداد بیمه با مشتریان بدین صورت است: هر مشترک به صورت مستقل در طی سال با احتمال p، یک میلیون تومان هزینه روی دست شرکت می‌گذارد ولی با احتمال $1-p$ هیچ خرجی برای شرکت نخواهد داشت. از طرفی، هر مشتری در طی سال حق بیمه ثابت و غیرتصادفی $1.003 \times p$ میلیون تومان را به شرکت پرداخت می‌کند.

الف) میانگین و انحراف معیار (جذر واریانس) کل هزینه سالانه شرکت چقدر است؟
ب) فرض کنید n برابر با یک میلیون نفر، و $p = 1/2$ باشد. یک کران بالا برای احتمال اینکه شرکت در آخر سال **ضرر کند** بدست بیاورید.

راهنمایی ۱: از تقریب قضیه حد مرکزی استفاده کنید. همچنین می‌دانیم که اگر $Z \sim \mathcal{N}(0,1)$، آنگاه برای هر $t > 0$ داریم $\mathbb{P}(Z > t) \leq e^{-t^2/2}$.
راهنمایی ۲: یکی از این اعداد شاید به دردتان بخورد:

x	۲/۰	۲/۵	۳/۰	۳/۵	۴/۰	۴/۵	۵/۰
e^{-x}	۰/۱۴	۰/۰۸	۰/۰۵	۰/۰۳	۰/۰۲	۰/۰۱	۰/۰۰۷

جواب:

الف) میزان هزینه‌ای که هر نفر در طی سال برای شرکت خواهد داشت را می‌توان به صورت X_i در نظر گرفت (برای $i = 1,\ldots,n$) که با حتمال p برابر با ۱ و با احتمال $1-p$ برابر با صفر است. پس توزیع برنولی دارد که می‌دانیم میانگین p و واریانسی برابر با $p(1-p)$ دارد.
کل هزینه شرکت برابر است با

$$C = X_1 + X_2 + \ldots + X_n$$

در نتیجه داریم $\mathbb{E}(C) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n) = np$، و با توجه به استقلال X_iها داریم:

$$\mathrm{Var}(C) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = np(1-p)$$

و لذا نتیجه می‌شود $\sigma_C = \sqrt{np(1-p)}$.
(می‌توانستیم از همان اول استدلال کنیم که توزیع C یک دوجمله‌ای با پارامترهای n, p است و روابط را مستقیم از حفظ بنویسیم).
پس جواب‌ها خواهد شد: میانگین np و انحراف معیار $\sqrt{np(1-p)}$.

ب) درآمد سالانه شرکت مقدار ثابت و غیرتصادفی $1.003 \times p \times n$ است. شرکت زمانی ضرر می‌کند که هزینه‌های آن از درآمدش بیشتر شود. در نتیجه:

$$\mathbb{P}(\text{loss}) = \mathbb{P}\left(X_1 + \ldots + X_n > 1.003np\right) \quad \Rightarrow$$

$$\mathbb{P}(\text{loss}) = \mathbb{P}\left(\frac{X_1 + \ldots + X_n - np}{\sqrt{np(1-p)}} > \frac{1.003np - np}{\sqrt{np(1-p)}}\right) = \mathbb{P}\left(Z_n > \frac{0.003p\sqrt{n}}{\sqrt{p(1-p)}}\right)$$

با توجه به اعداد داده شده داریم:

$$\frac{0.003p\sqrt{n}}{\sqrt{p(1-p)}} = 0.003 \times 1 \times 1000 = 3$$

از طرفی طبق تقریب **قضیه حد مرکزی** (با توجه به $n \gg 1$) می‌دانیم:

$$Z_n \triangleq \frac{X_1 + \ldots + X_n - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0,1)$$

لذا با فرض اینکه متغیر تصادفی $Z \sim \mathcal{N}(0,1)$ یک توزیع نرمال با میانگین صفر و واریانس ۱ داشته باشد، خواهیم داشت:

$$\mathbb{P}(\text{loss}) \simeq \mathbb{P}\left(Z \geq 3\right) \leq e^{-3^2/2} = e^{-4.5} \simeq 0.01$$

پس جواب نهایی ۱٪ خواهد بود.

سوال ۴ (۵ نمره):

فرض کنید متحرکی در نقطه صفر محور اعداد ایستاده است، و سپس به صورت گام به گام شروع به یک حرکت تصادفی (Random Walk) می‌کند. بدین صورت که در هر گام به طور مستقل، با احتمال‌های ۱/۲ یک واحد به <u>چپ</u> یا <u>راست</u> حرکت می‌کند. فرض کنید که این متحرک n گام تصادفی را برداشته باشد، به طوری که n عددی زوج است.

الف) استدلال کنید که امکان ندارد متحرک بعد از اتمام n حرکت روی یک نقطه فرد ایستاده باشد.

ب) k عددی زوج در بازه اعداد صحیح بین $-n$ تا n است. احتمال اینکه متحرک بعد از اتمام حرکت در نقطه k ایستاده باشد چقدر است؟

ج) داریم $n = 100$. نشان دهید احتمال اینکه متحرک در طی حرکت حداقل ۱۰ بار متوالی به راست حرکت کرده باشد از 0.09 کمتر است. **(راهنمایی) از کران اجتماع استفاده کنید:** $\mathbb{P}\left(A_1 \cup \ldots \cup A_m\right) \leq \mathbb{P}\left(A_1\right) + \ldots + \mathbb{P}\left(A_m\right)$.

جواب:

الف) فرض کنید متحرک در مجموع R بار به سمت راست، و L بار به سمت چپ حرکت کرده باشد (ترتیب‌ها مهم نیستند، خود تعداد دفعات اهمیت دارند). به وضوح داریم $L = n - R$. از طرفی، جایی که متحرک بعد از اتمام حرکت در آنجا ایستاده برابر است با $R - L$ (تعداد گام‌های حرکت به سمت راست منهای تعداد گام‌های حرکت کرده به سمت چپ)، که به عبارتی می‌شود $2R - n$، که حتماً عددی زوج خواهد بود. پس احتمال توقف نهایی در نقطه‌ای فرد برابر با صفر است.

ب) فرض کنید داشته باشیم $R - L = k$ که عددی زوج در بازه $-n$ تا n است. از طرفی داریم $R + L = n$. در نتیجه خواهیم داشت:

$$R = (n+k)/2 \quad , \quad L = (n-k)/2$$

که حتماً اعداد صحیح نامنفی خواهند شد. لذا با توجه به مفهوم توزیع دو جمله‌ای و احتمال‌های برابر (۱/۲) برای حرکت به سمت چپ یا راست، خواهیم داشت:

$$\mathbb{P}\left(\text{End point} = k\right) = \binom{n}{\frac{n+k}{2}}\frac{1}{2^n} = \binom{n}{\frac{n-k}{2}}\frac{1}{2^n}.$$

ج) واقعه A_i برای $i = 1,2,\ldots,91$ را تعریف می‌کنیم اینکه: متحرک از حرکت iام تا $(i+9)$ام (۱۰ گام متوالی) به سمت راست حرکت کرده باشد. احتمال اینکه متحرک زمانی در طی حرکت خود (هر زمانی) ۱۰ گام متوالی به سمت راست برداشته باشد برابر خواهد بود با:

$$\mathbb{P}\left(A_1 \cup \ldots \cup A_{91}\right)$$

که طبق راهنمایی سوال به صورت زیر کران‌دار می‌شود:

$$\mathbb{P}\left(A_1 \cup \ldots \cup A_{91}\right) \leq \mathbb{P}\left(A_1\right) + \ldots + \mathbb{P}\left(A_{91}\right)$$

از طرفی، احتمال هر یک از A_iها برابر با $(1/2)^{10}$ است، زیرا گام‌ها از یک دیگر مستقل بوده، و احتمال هر گام به راست ۱/۲ است. پس برای احتمال خواسته شده داریم:

$$\leq 91 \times \left(\frac{1}{2}\right)^{10} = \frac{91}{1024} < 0.09$$

سوال ۵ (۴ نمره):

یک کشور دوردست معمولاً در حالت خشکسالی به سرمی‌برد، اما به صورت کاملاً تصادفی در هر ۱۰ سال به طور متوسط یک سال پرباران را تجربه می‌کند. در سال‌های عادی توزیع بارش سالانه بر حسب میلی‌متر از یک توزیع پوآسون با پارامتر λ پیروی می‌کند. ولی در سالهای پرباران، پارامتر این توزیع پوآسون به 2λ افزایش می‌یابد.

الف) اگر یک سال را به صورت تصادفی انتخاب کنیم، احتمال اینکه در آن سال باران نبارد چقدر است؟

ب) در صورتیکه بدانیم در یک سال تصادفی، میزان بارش n میلی‌متر بوده است، احتمال اینکه آن سال یکی از سالهای پرباران باشد چقدر است؟

جواب:

الف) احتمال بارش صفر در سال خشک برابر است با

$$\mathbb{P}(X = 0 \,|\, \text{dry}) = e^{-\lambda}\frac{\lambda^0}{0!} = e^{-\lambda}\,.$$

احتمال بارش صفر در هنگام تَرسالی نیز برابر با

$$\mathbb{P}(X = 0 \,|\, \text{rainy}) = e^{-2\lambda}\frac{(2\lambda)^0}{0!} = e^{-2\lambda}\,.$$

در نتیجه، طبق قانون احتمال کل برای جواب نهایی قسمت الف) داریم:

$$\mathbb{P}(X = 0) = \frac{9}{10}\mathbb{P}(X = 0 \,|\, \text{dry}) + \frac{1}{10}\mathbb{P}(X = 0 \,|\, \text{rainy}) = \frac{9}{10}e^{-\lambda} + \frac{1}{10}e^{-2\lambda}$$

ب) راه حل این بخش از طریق قانون بیز است:

$$\mathbb{P}(\text{rainy} \,|\, n) = \frac{\mathbb{P}(n \,|\, \text{rainy})\mathbb{P}(\text{rainy})}{\mathbb{P}(n)} = \frac{\mathbb{P}(n \,|\, 2\lambda)\frac{1}{10}}{\mathbb{P}(n \,|\, 2\lambda)\frac{1}{10} + \mathbb{P}(n \,|\, \lambda)\frac{9}{10}}$$

(رابطه مخرج کسر فوق به دلیل «قانون احتمال کل» برقرار است).

احتمالات $\mathbb{P}(n\,|\,\lambda)$ و $\mathbb{P}(n\,|\,2\lambda)$ طبق فرض سوال پواسون هستند. از طرفی، داریم: $\mathbb{P}(\text{rainy}) = 1/10$. همچنین، داریم $\mathbb{P}(\text{dry}) = 9/10$، چرا که فرض شد سال مورد نظر به صورت «کاملاً تصادفی» انتخاب شده است لذا با احتمال ۱/۱۰ تَرسالی اتفاق افتاده است. در این صورت، جواب نهایی بدین صورت خواهد شد:

$$\mathbb{P}(\text{rainy}\,|\,n) = \frac{e^{-2\lambda}\frac{(2\lambda)^n}{n!}\cdot\frac{1}{10}}{e^{-2\lambda}\frac{(2\lambda)^n}{n!}\cdot\frac{1}{10} + e^{-\lambda}\frac{\lambda^n}{n!}\cdot\frac{9}{10}} = = \frac{1}{1+9e^{-(\lambda-2\lambda)}\left(\frac{\lambda}{2\lambda}\right)^n} = \frac{1}{1+9e^{\lambda}\left(\frac{1}{2}\right)^n}$$