دانشکده‌ی مهندسی کامپیوتر

آمار و احتمال مهندسی	نیم‌سال دوم ۱۴۰۱-۱۴۰۰

پاسخنامه تمرین سری سوم

مدرس: مهدی جعفری	موعد تحویل: پنج‌شنبه ۱۵ اردیبهشت ۱۴۰۱

سؤال ۱ برای n مرحله این کار را تکرار می‌کنیم که در نهایت به 5^n حالت مختلف ممکن است رخ دهد. حال X_{max} را متغیر تصادفی بیشینه در n بار آزمایش می‌نامیم:

$$\mathbb{P}(X_{max}=1)=\frac{1}{5^n}$$

$$\mathbb{P}(X_{max}=2)=\frac{2^n-1}{5^n}$$

$$\mathbb{P}(X_{max}=3)=\frac{3^n-2^n}{5^n}$$

$$\mathbb{P}(X_{max}=4)=\frac{4^n-3^n}{5^n}$$

$$\mathbb{P}(X_{max}=5)=\frac{5^n-4^n}{5^n}$$

حال امید ریاضی بیشینه اعداد یادداشت شده در این آزمایش را محاسبه می‌کنیم:

$$\mathbb{E}[X_{max}]=\sum_{i=1}^{5} i\mathbb{P}(x_{max}=i)=1\times\frac{1}{5^n}+2\times\frac{2^n-1}{5^n}+3\times\frac{3^n-2^n}{5^n}+4\times\frac{4^n-3^n}{5^n}+5\times\frac{5^n-4^n}{5^n}$$

$$\rightarrow \mathbb{E}[X_{max}]=\frac{5^{n+1}-4^n-3^n-2^n-1^n}{5^n}=5-\frac{4^n+3^n+2^n+1^n}{5^n}$$

سؤال ۲ الف) می‌داینم که آزمایش آخر قطعا موفق شده است، پس داریم:

$$\mathbb{P}(X=k)=C(k-1,m-1)p^m(1-p)^{k-m}$$

ب) با استفاده از اثبات لم زیر، موارد خواسته شده را به دست می‌آوریم:

اگر $X_1, X_2, ..., X_n$ متغیر های تصادفی مستقل با توزیع $Geometric(p)$ آنگاه متغیر تصادفی X که به صورت $X = X_1 + X_2 + ... + X_n$ است، دارای توزیع $Pascal(n, p)$ خواهد بود. زیرا داریم:

$$\mathbb{P}(X=k)=\mathbb{P}(X_1+...+X_n=k)$$

$$=\sum_{a_1+a_2+...+a_n=k}\mathbb{P}(X_1=a_1,...,X_n=a_n)$$

$$=\sum_{a_1+a_2+...+a_n=k}\mathbb{P}(X_1=a_1)\mathbb{P}(X_2=a_2)...\mathbb{P}(X_n=a_n)$$

$$= \sum_{a_1+a_2+...+a_n=k} p(1-p)^{a_1-1}p(1-p)^{a_2-1}...p(1-p)^{a_n-1}$$

$$= \sum_{a_1+a_2+...+a_n=k} p^n(1-p)^{k-n} = C(k-1,n-1)p^n(1-p)^{k-n}$$

پس خواهیم داشت:

$$X \sim Pascal(n,p)$$

حال با توجه به لمی که اثبات کردیم، خواهیم داشت:

$$Y = Y_1 + ... + Y_l, X = X_1 + ... + X_m$$

$$\Rightarrow Z = X_1 + ... + X_m + Y_1 + ... + Y_l$$

پس خواهیم داشت:

$$Z \sim Pascal(m+l,p)$$

می‌دانیمم که متغیر تصادفی‌ای که از توزیع هندسی پیروی می‌کند دارای امیدریاضی $\frac{1}{p}$ است. حال برای متغیر Z داریم:

$$\mathbb{E}[Z] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + ... + \mathbb{E}[X_m] + \mathbb{E}[Y_1] + \mathbb{E}[Y_2] + ... + \mathbb{E}[Y_l]$$

پس امیدریاضی متغیر تصادفی Z برابر جمع $m+l$ تا امیدریاضی متغیر تصادفی هندسی می‌باشد:

$$\mathbb{E}[Z] = \frac{m+l}{p}$$

سؤال ۳ الف) سوال را با فرض توزیع گسسته حل می‌کنیم: مقادیر برد توزیع گسسته را به صورت اعداد مرتب شده $x_1 < x_2 < ... < x_n$ نمایش می‌دهیم. حال فرض می‌کنیم که $x_i \leq c < x_{i+1}$ قرار دارد (برای معنادار بودن نمادها می‌توان در نظر گرفت که: $x_0 = -\infty$ و $x_{n+1} = \infty$). حال داریم:

$$f(c) = \mathbb{E}[|X-c|] = \sum_{x \leq x_i} -(x-c)\mathbb{P}[X=x] + \sum_{x > x_i} (x-c)\mathbb{P}[X=x]$$

$$= \sum_{x \leq x_i} (c-x)\mathbb{P}[X=x] + \sum_{x > x_i} (x-c)\mathbb{P}[X=x]$$

تابع بالا به ازای تمام cها پیوسته و در نتیجه مشتق‌پذیر است. برای نشان‌دادن مشتق‌پذیری در حالاتی که نابرابری معرفی شده برا c اکید نباشد بدیهی است. اگر c برابر یکی از x_iها شود نیز جمله $x-c$ در آخرین جمله جمع اول در حالت حدی مقدار صفر می‌گیرد و عبارت نهایی با مقدار $f(c)$ باز هم برابر است. پس در مجموع تابع به دست آمده مشتق‌پذیر است. با مشتق نسبت به c داریم:

$$\frac{\partial f(c)}{\partial c} = \sum_{x <= x_i} \mathbb{P}[X=x] - \sum_{x > x_i} \mathbb{P}[X=x]$$

$$= F(x_i) - (1 - F(x_i)) = 2F(x_i) - 1$$

که در آن $F(x)$ تابع توزیع تجمعی است. همچنین توجه داریم که چون $x_i \leq c < x_{i+1}$، بنابراین داریم: $F(x_i) = F(c)$.
با صفر قرار دادن مشتق به مقدار کمینه می‌رسیم (چون مشتق دوم نامنفی است، پس کمینه دارد):

$$2F(c) - 1 = 0 \Rightarrow F(c) = \frac{1}{2}$$

که عبارت به دست آمده در بالا همان تعریف میانه است. بنابراین در نقطه میانه، تابع مورد نظر مقدار کمینه خود را می‌گیرد.

ب) در واقع به دنبال آن هستیم که مقدار عبارت زیر را کمینه کنیم:

$$\mathbb{E}[(Y - f(X))^2|X = x]$$

حال دقت می‌کنیم که:

$$\mathbb{E}[(Y - f(X))^2|X = x] = \mathbb{E}[f(X)^2|X = x] + \mathbb{E}[Y^2|X = x] - 2\mathbb{E}[Yf(X)|X = x]$$
$$= f(x)^2 + \mathbb{E}[Y^2|X = x] - 2f(x)\mathbb{E}[Y|X = x]$$

با مشتق گرفتن نسبت به $f(x)$ و برابر صفر قرار دادن آن داریم:

$$2f(x) - 2\mathbb{E}[Y|X = x] = 0 \Rightarrow f(x) = \mathbb{E}[Y|X = x]$$

دقت داریم که با توجه به دانسته $X = x$ دو نمایش $f(X)$ و $f(x)$ معادل‌اند.

سؤال ۴ از خاصیت خطی بودن امید ریاضی روی احتمال شرطی استفاده می‌کنیم:

$$\mathbb{E}[X_1+X_2+...+X_n|X_1+X_2+...+X_n = x] = \mathbb{E}[X_1|X_1+X_2+...+X_n = x]+...+\mathbb{E}[X_n|X_1+X_2+...+X_n = x]$$

از طرفی با دانستن فرض $X_1 + X_2 + ... + X_n = x$ بدیهی است که داریم:

$$\mathbb{E}[X_1 + X_2 + ... + X_n|X_1 + X_2 + ... + X_n = x] = x$$

پس:

$$\mathbb{E}[X_1|X_1 + X_2 + ... + X_n = x] + ... + \mathbb{E}[X_n|X_1 + X_2 + ... + X_n = x] = x$$

حال با توجه به فرض iid بودن X_i ها داریم که:

$$\mathbb{E}[X_1|X_1 + X_2 + ... + X_n = x] = ... = \mathbb{E}[X_n|X_1 + X_2 + ... + X_n = x]$$

پس در نهایت:

$$n\mathbb{E}[X_1|X_1 + X_2 + ... + X_n = x] = x \Rightarrow \mathbb{E}[X_1|X_1 + X_2 + ... + X_n = x] = \frac{x}{n}$$

سؤال ۵ توجه: روابط نوشته شده مخصوص متغیرهای تصادفی گسسته در این سوال بدیهتا برای متغیرهای پیوسته نیز قابل تعمیم هستند.

الف)

$$\mathbb{E}[X|Y = y] = \sum_x x\mathbb{P}_{X|Y}(x|y) \quad (*)$$

از طرفی داریم:

$$X \perp Y \Rightarrow \mathbb{P}_{X|Y}(x|y) = \mathbb{P}_X(x)$$

که با جایگذاری در $(*)$ داریم:

$$\mathbb{E}[X|Y = y] = \sum_x x\mathbb{P}_X(x) = \mathbb{E}[X]$$

ب) توجه داریم که با توجه به شرط $Y = y$، می‌توان به جای متغیر تصادفی Y مقدار ثابت y را جایگذاری کرد:

$$\mathbb{E}[XY|Y = y] = \mathbb{E}[yX|Y = y] = y\mathbb{E}[X|Y = y]$$

توجه داریم که رابطه آخر طبق خاصیت خطی امید ریاضی به دست می‌آید که به شکل زیر به راحتی قابل تحقیق است (که برای احتمال شرطی نیز به دلیل مشابه برقرار است):

$$\mathbb{E}[aX] = \sum_x ax\mathbb{P}(x) = a\sum_x x\mathbb{P}(x) = a\mathbb{E}[X]$$

ج)

$$\mathbb{E}[Y\mathbb{E}[X|Y]] = \sum_y y\mathbb{E}[X|Y=y]\mathbb{P}(Y=y)$$

$$\overset{\text{طبق ب}}{=} \sum_y \mathbb{E}[XY|Y=y]\mathbb{P}(Y=y)$$

$$= \mathbb{E}_Y[\mathbb{E}[XY|Y]] \overset{\text{امید ریاضی کل}}{=} \mathbb{E}[XY]$$

سؤال ۶ اگر متغیر تصادفی X_k تعداد دفعاتی که باید لپ لپ خریداری کرد باشد، تا یک جایزه ی جدید ببینیم، درصورتی که قبلا k تا جایزه ی متفاوت دیده باشیم، آنگاه X تعداد دفعاتی است که باید لپ‌لپ بخریم تا از هر جایزه، یکی داشته باشیم. پس خواهیم داشت:

$$X = X_0 + X_1 + X_2 + ... + X_{n-1}$$

بنا به خاصیت خطی امیدریاضی داریم:

$$\mathbb{E}[X] = \mathbb{E}[X_0] + \mathbb{E}[X_1] + \mathbb{E}[X_2] + ... + \mathbb{E}[X_{n-1}]$$

از آنجایی که توزیع جوایز یکنواخت است، زمانی که، k تا جایزه ی متمایز داشته باشیم، احتمال آنکه جایزه ی جدید ببینیم برابر $\frac{n-k}{n}$ است. پس میانگین تعداد لپ‌لپ هایی که نیاز است بخریم تا جایزه‌ی جدید ببینیم برابر با: $\frac{n}{n-k}$ پس خواهیم داشت:

$$\mathbb{E}[X] = \frac{n}{n-0} + \frac{n}{n-1} + \frac{n}{n-2} + ... + \frac{n}{n-n+1}$$

پس در نهایت داریم:

$$\mathbb{E}[X] = n(\frac{1}{n} + \frac{1}{n-1} + ... + \frac{1}{1})$$

سؤال ۷ اگر فرض کنیم که 2^n تیم در این لیگ وجود دارد، در آن صورت n دور مسابقه در این لیگ داریم. در دور اول 2^{n-1} بازی برگزار می‌شود، در دور دوم 2^{n-2} تا دور n ام که ۱ بازی انجام می‌شود. پس تعداد کل بازی ها برابر $2^{n-1} + 2^{n-2} + ... + 2 + 1 = 2^n - 1$ است. حال یک بازی به خصوص مانند g را در نظر بگیرید. I_g متغیری است که مشخص می‌کند آیا فرد امتیاز این بازی را می‌گیرد یا نه. اگر در این بازی در دور $r = r(g)$ ام انجام بگیرد، میزان امتیازی که این فرد کسب خواهد کرد برابر 2^{r-1} خواهد بود. پس به طور متوسط امتیاز کسب شده از این بازی برابر است با:

$$Expected\ \ points\ \ g = E[2^{r-1}I_g] = 2^{r-1}E[I_g] = 2^{r-1}P_g$$

که P_g احتمال تشخیص برنده این بازی است. تشخیص درست برنده این بازی به منزله تشخیص درست نتیجه تمام بازی‌های قبلی آن تیم برنده در دور r ام است. پس $P_g = 2^{-r}$ و خواهیم داشت:

$$Expected\ \ points\ \ g = 2^{r-1}P_g = 2^{r-1}2^{-r} = \frac{1}{2}$$

که مستقل از بازی است. امید ریاضی امتیاز کسب شده از هر بازی برابر $\frac{1}{2}$ است. با توجه به این‌که $2^n - 1$ بازی در این لیگ داریم، امید ریاضی امتیاز فرد برابر با $\frac{2^n-1}{2}$ است. در این لیگ ۶۴ تیم موجود است. پس $n = 6$ می‌باشد. بنابراین خواهیم داشت:

$$Expected\ \ points = \frac{2^6 - 1}{2} = 31.5$$

سؤال ۸ مطابق راهنمایی مطرح شده در سوال ابتدا نامنفی بودن تابع g را نشان می‌دهیم. بنابراین داریم:

$$g(X,Y) = \mathbb{E}_{p_X}[\ln(\frac{p_X}{p_Y})] \Rightarrow -g(X,Y) = -\mathbb{E}_{p_X}[\ln(\frac{p_X}{p_Y})] = \mathbb{E}_{p_X}[-\ln(\frac{p_X}{p_Y})] = \mathbb{E}_{p_X}[\ln(\frac{p_Y}{p_X})] \quad (1)$$

حال توجه می‌کنیم که تابع ln یک تابع مقعر است. زیرا پیوسته و با مشتق دوم همواره منفی است:

$$\frac{\partial ln(x)}{\partial x} = \frac{1}{x} \Rightarrow \frac{\partial^2 ln(x)}{\partial x^2} = -\frac{1}{x^2} < 0 \Rightarrow \text{مقعر است } ln$$

حال با توجه به نامساوی *Jensen* داریم:

$$\mathbb{E}_{p_X}[\ln(\frac{p_Y}{p_X})] \leq ln(\mathbb{E}_{p_X}[\frac{p_Y}{p_X}]) = ln(\sum_{i=1:p_X(x_i)\neq 0}^{100} p_X(x_i)\frac{p_Y(x_i)}{p_X(x_i)}) = ln(\sum_{i=1:p_X(x_i)\neq 0}^{100} p_Y(x_i))$$

$$\leq ln(1) = 0 \overset{\text{طبق(۱)}}{\Rightarrow} -g(X,Y) \leq 0 \Rightarrow g(X,Y) \geq 0 \quad (2)$$

حال متغیر تصادفی Y را توزیع یکنواخت روی فضای نمونه مسئله درنظر می‌گیریم:

$$p_Y(x) = \frac{1}{100}$$

حال داریم:

$$g(X,Y) = \mathbb{E}_{p_X}[\ln(\frac{p_X}{p_Y})] = \sum_{i=1}^{100} p_X(x_i)\ln(\frac{p_X(x_i)}{p_Y(x_i)})$$

$$= \sum_{x=1}^{100} p_X(x)\ln(\frac{p_X(x)}{p_Y(x)}) = \sum_{i=1}^{100} p_X(x_i)\ln(p_X(x_i)) - \sum_{x=1}^{100} p_X(x_i)\ln(p_Y(x_i))$$

$$= -f(X) - \ln(\frac{1}{100})\sum_{x=1}^{100} p_X(x_i) = -f(X) + \ln(100) \overset{\text{طبق(۲)}}{\Rightarrow} -f(X) + \ln(100) \geq 0$$

$$\Rightarrow f(X) \leq \ln(100) = 4.605$$

این مقدار کران بالایی برای تابع $f(X)$ است و طبق روابط بالا به راحتی مشخص است که این کران به ازای توزیع یکنواخت اتفاق می‌افتد.

سؤال ۹ الف)

$$\sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{E}[X|A_i] = \sum_{i=1}^{n} \mathbb{P}(A_i)(\sum_{x} x\mathbb{P}(X=x|A_i))$$

$$= \sum_{i=1}^{n}\sum_{x} x\mathbb{P}(X=x|A_i)\mathbb{P}(A_i)$$

$$= \sum_{x} x\sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{P}(X=x|A_i)$$

$$= \sum_{x} x\mathbb{P}(X=x) = \mathbb{E}[X]$$

ب) در هر مرحله از این فرآیند دو اتفاق ممکن است بیفتد: دو دست یک فرد همدیگر را بگیرند. در این صورت به امیدریاضی تعداد حلقه ها یکی اضافه می‌شود و از تعداد نفرات حاضر در اتاق یکی کم می‌شود زیرا دو دست آن فرد یکدیگر را گرفته‌اند و کاربرد دیگری ندارد. حالت دوم زمانی اتفاق می افتد که یک شخص، دست شخص دیگر را بگیرد که در این حلقه‌ای ایجاد نمی‌شود ولی از تعداد افراد یکی کم می‌شود. زیرا در ابتدا ۴ دست موجود بود اما در بعد از این اتفاق دو دست از آن‌ها امکان استفاده دارند.

فرض کنید اگر در ابتدا n فرد درون اتفاق باشند، $f(n)$ امیدریاضی تعداد دورها است. اگر A حالت اول و B حالت دوم باشد آنگاه براساس بخش الف خواهیم داشت:

$$f(n) = \mathbb{E}[X] = \mathbb{E}[X|A]\mathbb{P}(A) + \mathbb{E}[X|B]\mathbb{P}(B)$$

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

می‌دانیم که در ابتدا نیز $2n$ دست آزاد داشتیم که در کل بین آن‌ها می‌توانیم به $\frac{2n(2n-1)}{2}$ انتخاب کنیم که n تای آن‌ها مربوط به وقتی‌است که دو دست یک فرد یکسان را انتخاب می‌کنیم. پس احتمال پیشامد A برابر با $\frac{n}{\frac{2n(2n-1)}{2}}$ است. که همان $\frac{1}{2n-1}$ می‌باشد. از طرفی $\mathbb{E}[X|A]$ یعنی امیدریاضی تعداد دورها به شرط این که در حرکت اول دو دست یک فرد را انتخاب کنیم، برابر امیدریاضی برای حالت $n-1$ شخص است به علاوه‌ی یک دوری که ایجاد شده‌است. پس داریم که $f(1) = 1$

در نهایت خواهیم داشت:

$$f(n) = \mathbb{E}[X] = \frac{1}{2n-1} + \frac{1}{2n-3} + ... + \frac{1}{3} + 1 = H_{2n} - \frac{H_n}{2}$$

سؤال ۱۰ توجه کنید که داریم:

$$510510 = 2 \times 3 \times 5 \times 7 \times 11 \times 13 \times 17$$

حال متغیر تصادفی X_i را به این شکل تعریف می‌کنیم که: «اگر با عدد i شروع کنیم، در هر مرحله عوامل اول مقسوم‌علیه انتخاب شده را بر عدد حاصل تقسیم می‌کنیم. در صورت نبودن عامل اول در آن مقسوم‌علیه یا امکان‌پذیر نبودن تقسیم، عدد آن مرحله را بر 1 تقسیم می‌کنیم. عدد باقیمانده بعد از 10 مرحله را X_i می‌نامیم.» حال اگر عوامل اول عدد 510510 را در نظر بگیریم مشخص است که داریم:

$$X_{510510} = X_2 X_3 X_5 X_7 X_{11} X_{13} X_{17}$$

حال اگر i عددی اول باشد، چون تنها دو مقسوم علیه i و ۱ را دارد، در مجموع در تنها یک حالت برابر با i می‌ماند (اگر هربار تقسیم بر ۱ شود) یا به ۱ می‌رسد. به عبارتی:

$$\mathbb{P}(X_i = i) = \frac{1}{1024}, \quad \mathbb{P}(X_i = 1) = \frac{1023}{1024}$$

$$\mathbb{E}[X_i] = \frac{1}{1024} \times i + \frac{1023}{1024} \times 1 = \frac{1023 + i}{1024}$$

حال $\mathbb{E}[X_{510510}]$ را تشکیل می‌دهیم:

$$\mathbb{E}[X_{510510}] = \mathbb{E}[X_2 X_3 X_5 X_7 X_{11} X_{13} X_{17}] \overset{iid}{=} \mathbb{E}[X_2]\mathbb{E}[X_3]\mathbb{E}[X_5]\mathbb{E}[X_7]\mathbb{E}[X_{11}]\mathbb{E}[X_{13}]\mathbb{E}[X_{17}]$$
$$= \frac{1025 \times 1026 \times 1028 \times 1030 \times 1034 \times 1036 \times 1040}{1024^7} = 1.0508$$

موفق باشید