دانشکده‌ی مهندسی کامپیوتر

آمار و احتمال مهندسی | زمستان ۱۴۰۲

تمرین سری دوم

مدرس: دکتر مهدی جعفری | موعد تحویل: چهارشنبه ۲۳ اسفند ۱۴۰۲

سؤال ۱ پیامی بر یک کانال دارای نویز ارسال شده است. این پیام شامل n بیت $x_1, x_2, ..., x_n$ می‌باشد به طوری که $x_i \in \{0,1\}$. از آنجایی که کانال دارای نویز می‌باشد، شانسی برای تغییر هر یک از بیت‌ها و خراب شدن پیام وجود دارد که منجر به تولید خطا می‌شود (یعنی تبدیل یک به صفر یا برعکس). فرض کنید رخداد خطاها از یکدیگر مستقل هستند و p احتمال این است که یک بیت دارای خطا باشد به طوری که $0 < p < 0.5$. حال دنباله $y_1, y_2, ..., y_n$ را پیامی در نظر بگیرید که در نهایت به دست ما رسیده است (در این صورت اگر $y_i = x_i$ باشد یعنی خطایی رخ نداده است).
برای یافتن راحت‌تر خطا، بیت n ام را برای بررسی کردن زوجیت در نظر می‌گیریم. بدان معنا که x_n را برابر با صفر در نظر می‌گیریم اگر $x_1 + x_2 + ... + x_{n-1}$ زوج باشد و آن را برابر با یک تعریف می‌کنیم اگر مجموع گفته شده برابر با مقداری فرد باشد. هنگامی که پیام دریافت شد، بررسی می‌کنیم که y_n زوجیت برابری با $y_1 + y_2 + ... + y_{n-1}$ دارد یا خیر. اگر زوجیت یکسان نبود، متوجه می‌شویم که حداقل یک بیت دچار خطا شده است و در غیر این صورت خطایی وجود ندارد.
الف. برای $n = 5$ و $p = 0.1$، احتمال آن که پیام دریافتی دارای خطا(هایی) است و نتوانسته‌ایم آن(ها) را تشخیص دهیم مشخص کنید.
ب. برای حالت کلی n و p عبارتی (به صورت جمعی) بنویسید که بتواند قسمت الف را محاسبه کند.
ج. یک عبارت ساده شده که شامل جمع تعداد زیادی جمله نمی‌باشد برای قسمت‌های قبل و به صورت عمومی n و p بنویسید.

پاسخ:

الف.

توجه کنید $\sum_{i=1}^{n} x_i$ زوج می‌باشد. اگر تعداد خطاها زوج (و ناصفر) باشند، خطاها کشف نمی‌شوند. در غیر این صورت $\sum_{i=1}^{n} y_i$ فرد خواهد شد و خطاها کشف می‌شوند. تعداد خطاها برابر است با $Bin(n,p)$، در نتیجه احتمال خطاهای کشف نشده وقتی $n = 5$، $p = 0.1$ برابر است با:

$$\binom{5}{2} p^2 (1-p)^3 + \binom{5}{4} p^4 (1-p) \approx 0.073$$

ب.

با استدلالی مشابه در قسمت الف، احتمال خطاهای کشف نشده برابر است با:

$$\sum_{\substack{k \text{ even} \\ k \geq 2}} \binom{n}{k} p^k (1-p)^{n-k}$$

ج.

a، b را به صورت زیر تعریف می‌کنیم:

$$a = \sum_{\substack{k \text{ even} \\ k \geq 0}} \binom{n}{k} p^k (1-p)^{n-k}$$

$$b = \sum_{\substack{k \text{ odd} \\ k \geq 1}} \binom{n}{k} p^k (1-p)^{n-k}$$

در این صورت خواهیم داشت:

$$a + b = \sum_{k \geq 0} \binom{n}{k} p^k (1-p)^{n-k} = 1$$

$$a - b = \sum_{k \geq 0} \binom{n}{k} (-p)^k (1-p)^{n-k} = (1-2p)^n$$

با توجه به a و b این دستگاه را حل می‌کنیم:

$$a = \frac{1 + (1-2p)^n}{2}$$

$$b = \frac{1 - (1-2p)^n}{2}$$

در نتیجه داریم:

$$\sum_{\substack{k \text{ even} \\ k \geq 0}} \binom{n}{k} p^k (1-p)^{n-k} = \frac{1 + (1-2p)^n}{2}$$

حال احتمال اینکه خطایی نداشته باشیم را از این مقدار کم می‌کنیم:

$$\sum_{\substack{k \text{ even} \\ k \geq 2}} \binom{n}{k} p^k (1-p)^{n-k} = \frac{1 + (1-2p)^n}{2} - (1-p)^n$$

سؤال ۲ امروز توانسته ایم اولین کتاب آمار و احتمال را به کمک همدیگر چاپ کنیم. اگر در این کتاب، تعداد اشتباهات چاپی در هر صفحه را بتوان با یک متغیر تصادفی پوآسون و با پارامتر t مدل کرد و همچنین تعداد این اشتباهات در صفحه‌های مختلف کتاب از یکدیگر مستقل باشند،

الف. احتمال پیش‌آمد این که دومین اشتباه در صفحه k رخ داده باشد را محاسبه کنید.

ب. فرض کنید ویراستار این کتاب، هر کدام از اشتباهات چاپی را به احتمال p و مستقل از سایر اشتباهات بیابد. اگر X تعداد اشتباهاتی باشد که او متوجه آن‌ها شده است و Y تعداد اشتباهات باقی‌مانده باشد، آنگاه توزیع این دو متغیر تصادفی را بیابید و نشان دهید از هم مستقل هستند.

پاسخ:

الف.

پیشامد A را بگیرید آنکه دومین اشتباه در صفحه r رخ داده باشد و پیشامد B را نیز آن بگیرید که اشتباه اول در $r-1$ صفحه اول رخ داده باشد، آنگاه با توجه به قانون احتمال کل داریم:

$$P(A) = P(A|B) + P(A|B') \quad (1)$$

برای محاسبه $P(A|B)$ توجه کنید که در $r-1$ صفحه اول، باید دقیقا در یک صفحه فقط یک اشتباه رخ داده باشد و به علاوه در صفحه r ام بیش از یک خطا رخ داده باشد (مکمل پیشامد رخ ندادن هیچ خطایی)، که از مستقل بودن خطاها در صفحه‌های متفاوت خواهیم داشت:

$$P(A|B) = \binom{r-1}{1} t\frac{e^t}{1!}\left(\frac{e^t}{0!}\right)^{r-2}\left(1-\frac{e^t}{0!}\right) = rte^{t(r-1)}(1-e^t) = rte^{t(r-1)} - rte^{rt}$$

برای محاسبه $P(A|B')$ نیز باید خطای اول و دوم در صفحه r رخ داده باشند و یا به عبارت دیگر در هیچ یک از $r-1$ صفحه اول خطایی رخ نداده باشد و در صفحه r ام بیش از دو خطا رخ دهد:

$$P(A|B') = \left(\frac{e^t}{0!}\right)^{r-1}\left(1-\frac{e^t}{0!}-t\frac{e^t}{1!}\right) = e^{t(r-1)} - (1+t)e^{rt}$$

حال از (۱) خواهیم داشت:

$$P(A) = rte^{t(r-1)} - rte^{rt} + e^{t(r-1)} - (1+t)e^{rt} = (rt+1)e^{t(r-1)} - (1+t+rt)e^{rt}$$

ب.

متغیر تصادفی N را تعداد کل اشتباهات در نظر بگیرید:

$$N \sim Poisson(t)$$

در این صورت توزیع X به شرط $N=n$ یک توزیع دوجمله‌ای با پارامترهای (n,p) است:

$$P(X=k|N=n) = \begin{cases} \binom{n}{k}p^k(1-p)^{n-k} & k \le n \\ 0 & k > n \end{cases} \qquad (2)$$

حال از قانون احتمال کل داریم:

$$P(X=k) = \sum_{n=0}^{\infty} P(X=k, N=n) = P(N=n)P(X=k|N=n)$$

که مقدار $P(X=k|N=n)$ برای مقادیر $k>n$ صفر است. در ادامه با استفاده از (۱) داریم:

$$\begin{aligned} P(X=k) &= \sum_{n=k}^{\infty} e^{-t}\frac{t^n}{n!}\binom{n}{k}p^k(1-p)^{n-k} \\ &= \frac{e^{-t}}{k!}p^k\sum_{n=k}^{\infty}\frac{t^n}{(n-k)!}(1-p)^{n-k} \\ &= \frac{e^{-t}}{k!}p^k\sum_{s=0}^{\infty}\frac{t^{k+s}}{s!}(1-p)^s \quad (s := n-k) \\ &= \frac{(tp)^k}{k!}e^t\sum_{s=0}^{\infty}\left(\frac{(t(1-p))^s}{s!}\right) \\ &= \frac{(tp)^k}{k!}e^{-t}e^{t(1-p)} \\ &= \frac{(tp)^k}{k!}e^{-tp} \end{aligned}$$

که از تساوی بالا نتیجه می‌شود متغیر تصادفی X، توزیع پوآسون با پارامتر tp دارد. همینطور به صورت مشابه نتیجه می‌شود متغیر Y توزیع پوآسون با پارامتر $t(1-p)$ دارد. اکنون برای نشان دادن مستقل بودن این دو کافی است تعریف استقلال را بنویسیم:

$$P(X=k, Y=s) = P(X=k)P(Y=s)$$

برای این منظور داریم:

$$P(X=k, Y=s) = P(X=k)P(N=k+s)$$

$$
\begin{aligned}
&= P(N=k+s)P(X=k|N=k+s) \\
&= e^{-t}\frac{t^{k+s}}{(k+s)!}\binom{k+s}{k}p^k(1-p)^s \\
&= e^{-tk}e^{-ts}\frac{t^k t^s}{k!s!}p^k(1-p)^s \\
&= \left(e^{-tk}\frac{(tp)^k}{k!}\right)\left(e^{-st}\frac{((1-p)t)^s}{s!}\right) \\
&= P_X(k)P_Y(s)
\end{aligned}
$$

سؤال ۳ فرض کنید متحرکی در مبدأ مختصات قرار دارد و سپس به صورت گام به گام شروع به یک حرکت تصادفی (*Random Walk*) می‌کند. بدین صورت که در هر گام، مستقل از گام‌های پیشین، با احتمال‌های برابر 0.25 یک واحد به بالا، پایین، چپ و یا راست حرکت می‌کند. فرض کنید این متحرک k گام تصادفی را برداشته باشد.

الف. احتمال اینکه متحرک در این لحظه، در مختصات (m, n) ایستاده باشد چقدر است؟ (با فرض اینکه $m, n \in \mathbb{Z}$ و $|m| + |n| \leq k$).

ب. احتمال اینکه متحرک صرفاً در نقطه‌ای با $x = m$ (بدون هیچ‌گونه شرط بر y) ایستاده باشد چقدر است؟

ج. در صورتی که بدانیم از k گام برداشته شده، دقیقاً تعداد k' (که $k' \leq k$) گام در راستای محور x، یعنی چپ یا راست بوده‌اند، سوال قسمت ب را مجدد پاسخ دهید.

پاسخ:

الف.

فرض کنید متحرک در طی k گامی که برداشته، مقدار (r, l, u, d) گام را به ترتیب به سمت راست، چپ، بالا و پایین بردارد. همواره داریم $r + l + u + d = k$. چون قرار است در مختصه $x = m$ ایستاده باشد باید داشته باشیم: $r - l = m$، یا معادلا $r = r + m$. همچنین چون قرار است در مختصه $y = n$ نیز باشیم، داریم $u = d + n$. دقت کنید با توجه به استنتاج‌های بالا، اگر $k - |m| - |n|$ عددی زوج نشد، احتمال رسیدن به نقطه (m, n) در k گام، صفر است. در نتیجه، برای احتمال رسیدن به نقطه (m, n) می‌توان از توزیع چندجمله‌ای استفاده کرد و خواهیم داشت:

$$
P_{(m,n)} = \sum_{(r,l,u,d)\in S} \frac{k!}{r!l!u!d!}\left(\frac{1}{4}\right)^k
$$

که برای مجموعه S داریم:

$$
S = \{(r,l,u,d) \mid r,l,u,d \geq 0,\ r+l+u+d = k,\ r-l = m,\ u-d = n\}
$$

نکته‌ای که ممکن است شما به آن توجه کرده باشید این است که به ازای مقادیر دلخواه از k, m, n برخی از نقاط صفحه دارای $P_{(m,n)} = 0$ هستند. یعنی به ازای نقطه‌ای مشخص مانند (m, n)، با برخی مقادیر k امکان ندارد بتوان به آن‌ها رسید.

ب.

بدون کاستن از کلیت مسئله فرض کنید $m \geq 0$ باشد. در این صورت، می‌توان یک توزیع چندجمله‌ای با احتمال‌های $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right)$ برای حرکت به راست، چپ و یا حرکت عمودی در نظر گرفت. در نتیجه، می‌توان نوشت:

$$
P_m = \sum_{l=0}^{\lfloor\frac{k-m}{2}\rfloor} \frac{k!}{l!(l+m)!(k-2l-m)!}\left(\frac{1}{4}\right)^{2l+m}\left(\frac{1}{2}\right)^{k-2l-m}
$$

همچنین به راحتی می‌توان استنتاج کرد که $P_m = P_{-m}$. پس می‌توانید در رابطه بالا m را به $|m|$ تغییر دهید.

ج.

در صورت دانستن k' جواب بسیار ساده‌تر می‌شود (دقت کنید که مسئله کاملا یک بعدی خواهد شد و احتمال حرکت به چپ یا راست، هرکدام $\frac{1}{2}$ خواهد شد):

$$P_{m|k'} = \binom{\frac{k'-|m|}{2}}{k'} \left(\frac{1}{2}\right)^{\frac{k'-|m|}{2}} \left(\frac{1}{2}\right)^{\frac{k'+|m|}{2}} = \binom{\frac{k'-|m|}{2}}{k'} \left(\frac{1}{2}\right)^{k'}$$

در صورتی که $k' - m$ عددی زوج نشد، $P_{(m|k')}$ برابر با صفر است.

موفق باشید