بسمه تعالی

آمار و احتمال مهندسی

نیمسال دوم ۱۴۰۱-۱۴۰۲

مدرس: الهام منیفی

دانشکده مهندسی کامپیوتر

تمرین پنجم

سوال اول

نامساوی چبی‌شف:

$$P[|X - E[X]| \geq a] \leq \frac{Var[X]}{a^2}$$

باتوجه به اینکه $E[\overline{X}_n] = \frac{1}{n}\sum_{i=1}^{n} \mu_i$ و $Var(\overline{X}_n) = \frac{1}{n}\sum_{i=1}^{n^2} Var(X_i)$ با جای‌گذاری این عبارت در نامساوی چبی‌شف به نامساوی می‌رسیم.

بخش ب)

باتوجه به کران‌دار بودن $Var(X_i) \leq M$ داریم:

$$N \to \infty : Var(\overline{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) \leq \frac{1}{n}M \to 0$$

پاسخ سوال دوم در ابتدا این معادله به ازای $n = 0$ درست است (صحت فرض پایه استقرا). در ادامه اگر فرض کنیم که این معادله به ازای n درست است:

$$P\{N(x) \geq n+1\} = \frac{x^{n-1}}{(n-1)!}$$

حال برای حالت $n+1$ خواهیم داشت:

$$P\{N(x) \geq n+1\} = \int_0^1 P\{N(x) \geq n | U_1 = y\} dy$$

$$= \int_0^x P\{N(x-y) \geq n\} dy$$

$$\int_0^x N(u) du$$

$$= \int_0^x \frac{u^{n-1}}{(n-1)!} du = \frac{x^n}{n!}$$

بخش ب

$$E[N(x)] = \sum_{n=۰}^{\infty} P\{N(x) \geq n\} = \sum_{n=۰}^{\infty} P\{N(x) \geq n + ۱\} = \sum_{n=۰}^{\infty} \frac{x^n}{n!} = e^x$$

Problem 3. (10 pts.) **(a)** Define

$$X_i = \begin{cases} 1 & \text{if person } i \text{ supports Erika} \\ 0 & \text{if person } i \text{ does not support Erika} \end{cases}$$

Then $X_i \sim \text{Bern}(0.5)$ and the number of people who prefer Erika is

$$S = X_1 + \cdots + X_{400}.$$

We know $E(X_i) = 1/2$ and $\text{Var}(X_i) = 1/4$. This implies $E(S) = 200$ and $\text{Var}(S) = 100$. Thus the central limit theorem tells us that

$$S \approx \text{N}(200, 100).$$

The problem asks for $P(S > 210)$:

$$P(S > 210) = P\left(\frac{S - 200}{10} > \frac{210 - 200}{10}\right) \approx P(Z > 1) \approx \boxed{0.16}.$$

(b) If we now let $Y_i = 1$ if person i prefers one of Peter, Jon or Jerry and 0 otherwise, we have $Y_1, \ldots, Y_{400}$ independent Bern(0.3). So $E(Y_i) = \mu = 0.3$ and $\text{Var}(Y_i) = (0.3)(0.7) = 0.21$. If $\overline{Y}_{400} = \frac{1}{400}(Y_1 + \cdots + Y_{400})$, the Central Limit Theorem tells us

$$\frac{\overline{Y}_n - \mu}{\sigma/\sqrt{400}} = \frac{(\overline{Y}_n - 0.3)\sqrt{400}}{\sqrt{0.21}}$$

is approximately standard normal. If Z is standard normal, then

$$P(\overline{Y} \leq 0.25) \approx P\left(Z < \frac{(0.25 - 0.3)\sqrt{400}}{\sqrt{0.21}}\right) \approx 0.0145.$$

سوال چهارم (نمره)

Solution. (a) We have

$$\begin{aligned} p_{X,Y,Z}(x,y,z) &= \mathbf{P}(X=x, Y=y, Z=z) \\ &= \mathbf{P}(X=x)\mathbf{P}(Y=y, Z=z \mid X=x) \\ &= \mathbf{P}(X=x)\mathbf{P}(Y=y \mid X=x)\mathbf{P}(Z=z \mid X=x, Y=y) \\ &= p_X(x)p_{Y|X}(y \mid x)p_{Z|X,Y}(z \mid x,y). \end{aligned}$$

(b) The formula can be written as

$$\mathbf{P}(X=x, Y=y, Z=z) = \mathbf{P}(X=x)\mathbf{P}(Y=y \mid X=x)\mathbf{P}(Z=z \mid X=x, Y=y),$$

which is a special case of the multiplication rule.

(c) The generalization is

$$\begin{aligned} &p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \\ &\qquad = p_{X_1}(x_1)p_{X_2|X_1}(x_2 \mid x_1)\cdots p_{X_n|X_1,\ldots,X_{n-1}}(x_n \mid x_1,\ldots,x_{n-1}). \end{aligned}$$

سوال پنجم (نمره)

اگر فرض کنیم متغیر تصادفی X به هر مبلغ، خطای آن را نسبت می‌دهد، آنگاه:

$$Y = \frac{\sum^{n} X_i}{n} \Rightarrow P = \mathbb{P}[\,|nY| \geq 1000\,] \overset{n=1000}{=\!=} \mathbb{P}[\,|Y| \geq 1\,]$$

طبق قضیه حد مرکزی : $Y \sim N(\mu_y, \sigma_y^2)$: $\mu_y = E\left[\frac{\sum X_i}{n}\right] = \frac{1}{n}(n\,E[X]) = E[X] = \mu_x$

$$, \ \sigma_y^2 = E[(Y-\mu_x)^2] = E\left[\left(\frac{\sum (X_i - \mu_x)}{n}\right)^2\right] = \frac{1}{n^2}(\sigma_x^2 \times n) = \frac{\sigma_x^2}{n}$$

$\Rightarrow Y \sim N\left(\mu_x, \frac{\sigma_x^2}{n}\right) \Rightarrow$ Y : خطای مبلغ می‌تواند از بازه $(-25, 25)$ باشد و وزن همه نقاط نیز یکسان فرض می‌شود، پس در واقع X، یونیفرم در بازه $(-25, 25)$ است، پس

$$E[X] = \frac{-25+25}{2} = 0 \ , \ \sigma_x^2 = E[X^2] - 0 = \int_{-25}^{25} x^2 \times \frac{1}{50}\,dx = \frac{1}{50}\left(\frac{1}{3}x^3\right)\Big|_{-25}^{25} = \frac{2 \times 25^3}{3 \times 50} = \frac{625}{3}$$

$n = 1000$

$$\Rightarrow Y \sim N\left(0, \frac{625}{3000}\right) \Rightarrow Y \sim N\left(0, \frac{5}{24}\right) \Rightarrow f_Y(y) = \frac{1}{\sqrt{\frac{5\pi}{12}}}\, e^{-\frac{12y^2}{5}}$$

$$\Rightarrow \mathbb{P}[\,|Y| \geq 1\,] = \mathbb{P}[Y \geq 1] + \mathbb{P}[Y \leq -1] = 1 - \mathbb{P}[Y \leq 1] + \mathbb{P}[Y \leq -1]$$

$$= 1 - F_Y(1) + F_Y(-1) = 1 - (F_Y(1) - F_Y(-1)) = 1 - \int_{-1}^{1} \frac{1}{\sqrt{\frac{5\pi}{12}}}\, e^{-\frac{12y^2}{5}}\,dy \simeq 1 - 0.014$$

$$\Rightarrow \mathbb{P}[\,|Y| \geq 1\,] \simeq \boxed{0.985}$$

سوال ششم

در ابتدا باید ماتریس ژاکوبین را تشکیل بدهیم.

$$Y_n = X_۱ + X_۲ + ... + X_n$$

در این صورت داریم:

$$J = \begin{bmatrix} ۱ & ۰ & ۰ & ... & ۰ \\ ۱ & ۱ & ۰ & ... & ۰ \\ ۱ & ۱ & ۱ & ... & ۰ \\ ... & & ... & & \\ ... & & ... & & \\ ۱ & ۱ & ۱ & ... & ۱ \end{bmatrix}$$

که دترمینان این ماتریس برابر یک خواهد بود. حال تنها کافی است که با بدست آوردن توزیع توام $f_{X_۱,X_۲,...,X_n}(x_۱, x_۲, ..., x_n)$ و تغییر متغیر آن به توزیع توام مد نظر برسیم:

$$f_{X_۱,X_۲,...,X_[}n(x_۱, x_۲, ..., x_n) = \Pi_{i=۱}^{n} \lambda e^{-\lambda x_i}$$

حال با انجام تغییر متغیر:

$$X_۱ = Y_۱, X_i = Y_i - Y_{i-۱}, i = ۱, ۲, ..., n$$

در نتیجه با جای‌گذاری داریم:

$$f_{Y_۱,Y_۲,...,Y_n}(y_۱, y_۲, ..., y_n) =$$
$$f_{X_۱,X_۲,...,X_n}(y_۱, y_۲ - y_۱, ..., y_n - y_{n-۱}) =$$
$$\lambda^n exp\{-\lambda[-y_۱ + \sum_{i=۲}^{n}(y_i - y_{i-۱}]\} = \lambda^n e^{-\lambda y_n}, ۰ < y_۱ < y_۲ < ... < y_n$$

بخش ب) حال برای بدست آوردن توزیع حاشیه‌ای Y_n بر روی متغیرهای $Y_۲, Y_۳, ..., Y_n$ انتگرال بگیریم. برای اینکار باتوجه به بازه‌های متغیرها که به صورت $۰ < y_۱ < y_۲ < ... < y_n$ است، از $y_۲$ شروع می‌کنیم:

$$f_{Y_۲,Y_۳,...,Y_n}(y_۲, y_۳, ..., y_n) = \int_۰^{y_۲} \lambda^n e^{-\lambda y_n} dy_۱ = \lambda^n y_۲ e^{-\lambda y_n}, ۰ < y_۲, y_۳ < ... < y_n$$

با ادامه دادن همین روند بر متغیر $y_۲$:

$$f_{Y_۳,Y_۴,...,Y_n}(y_۳, y_۴, ..., y_n) = \int_۰^{y_۳} \lambda^n y_۲ e^{-\lambda y_n} dy_۲ = \lambda^n \frac{y_۳^۲}{۲} e^{-\lambda y_n}, ۰ < y_۳ < y_۴ < ... < y_n$$

به همین ترتیب با انجام انتگرال بعدی:

$$f_{Y_۴,Y_۵,...,Y_n}(y_۴, y_۵, ..., y_n) = \lambda^n \frac{y_۴^۳}{۳!} e^{-y_n}, ۰ < y_۴ < y_۵ < ... < y_n$$

باتوجه به الگوی بدست‌آمده، درنهایت تابع توزیع حاشیه‌ای Y_n برابر است با:

$$f_{Y_n}(y_n) = \lambda^n \frac{y_n^{n-۱}}{(n-۱)!} e^{-\lambda y_n}, < y_n$$

موفق باشید.