محمد پیام تابش ۴۰۰۱۰۴۸۶۷ تمرین ۴

1)

$$cor(x, Y|Z) = E\left[(x - E[x|Z])(Y - E[Y|Z])|Z\right]$$

الف)

$$cor(x, Y|Z) = E\left[(x - E[x|Z])(Y - E[Y|Z])|Z\right]$$

$$= E\left[xY - xE[Y|Z] - YE[X|Z] + E[X|Z]E[Y|Z] \,\middle|\, Z\right]$$

$$= E[XY|Z] - E\left[xE[Y|Z]|Z\right] - E\left[YE[X|Z]|Z\right]$$

$$+ E\left[E[X|Z]E[Y|Z]|Z\right]$$

$$= E[XY|Z] - E[X|Z]E[Y|Z] - E[Y|Z]E[X|Z]$$

$$+ E[X|Z]E[Y|Z]$$

$$= E_{[XY|Z]} - E_{[X|Z]} E_{[Y|Z]}$$

$$\Rightarrow \operatorname{cov}(X,Y|Z) = E_{[XY|Z]} - E_{[X|Z]} E_{[Y|Z]}$$

ب)

$$\operatorname{cov}(X,Y) = E_{[XY]} - E_{[X]} E_{[Y]} \quad , \quad \boxed{E_{[E[\varphi|p]]} = E_{[\varphi]}}$$

$$= E_{[\underbrace{E[XY|Z]}]} - E_{[E[X|Z]]} E_{[E[Y|Z]]}$$

طبق الف

$$\boxed{E_{[XY|Z]} = \operatorname{cov}(X,Y|Z) + E_{[X|Z]} E_{[Y|Z]}}$$

$$\Rightarrow \operatorname{cov}(X,Y) =$$

$$E_{[\operatorname{cov}(X,Y|Z) + E_{[X|Z]} E_{[Y|Z]}]} - E_{[E[X|Z]]} E_{[E[Y|Z]]}$$

$$= E_{[\operatorname{cov}(X,Y|Z)]} + E_{[\underbrace{E_{[X|Z]}}_{U} \underbrace{E_{[Y|Z]}}_{V}]} - E_{[\underbrace{E[X|Z]}_{U}]} E_{[\underbrace{E[Y|Z]}_{V}]}$$

$$* \; E_{[UV]} - E_{[U]} E_{[V]} = \operatorname{cov}(U,V) \; *$$

$$\Rightarrow \boxed{\text{cov}(X,Y) = E[\text{cov}(X,Y|Z)] + \text{cov}\left(E[X|Z], E[Y|Z]\right)}$$

ج

$$\text{var}(X|Y) = E[(X - E[X|Y])^2 | Y]$$

$$= E[X^2 - 2XE[X|Y] + E^2[X|Y] \,|\, Y]$$

$$= E[X^2|Y] - 2E^2[X|Y] + E^2[X|Y]$$

$$\Rightarrow \text{var}(X|Y) = E[X^2|Y] - (E[X|Y])^2$$

$$\Rightarrow E[\text{var}(X|Y)] = E[E[X^2|Y]] - E[(E[X|Y])^2]$$

$$= E[X^2] - E[(E[X|Y])^2]$$ ⊕ رابطه‌ی

و از طرفی :

$$\text{var}(E[X|Y]) = E[(E[X|Y])^2] - (\underbrace{E[E[X|Y]]}_{= E[X]})^2$$

$$\Rightarrow var(E[X|Y]) = E[(E[X|Y])^2] - (E[X])^2$$ ★ رابطه

اگر ★ را با ⊕ جمع کنیم :

$$Var(E[X|Y]) + E(var(X|Y)) = E[X^2] - (E[X])^2$$

$$= Var(X)$$

$$\Rightarrow Var(X) = var(E[X|Y]) + E[var(X|Y)]$$

۲

$$Y = \sum_{i=1}^{۵۰} X_i = X_۱ + X_۲ + \cdots + X_{۵۰}$$

$$E[X_i] = ۲, \quad Var[X_i] = ۱$$

$$P(۹۰ < Y < ۱۱۰) = ؟$$

$$Z = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \qquad Z \sim Normal(۰,۱)$$

$n = ۵۰$

$$Z = \frac{Y - ۵۰ \times ۲}{۱\sqrt{۵۰}} = \frac{Y - ۱۰۰}{۵\sqrt{۲}}$$

$$P(۹۰ < Y < ۱۱۰) = P\left(\frac{۹۰-۱۰۰}{۵\sqrt{۲}} < Z < \frac{۱۱۰-۱۰۰}{۵\sqrt{۲}}\right)$$

$$= P(-\sqrt{۲} < Z < \sqrt{۲}) \quad , Z \sim N(۰,۱)$$

$$= \Phi(\sqrt{۲}) - \Phi(-\sqrt{۲})$$

$$= ۲\Phi(\sqrt{۲}) - ۱ = ۰/۸۴۲۷$$

$$\rightarrow \boxed{P(۹۰ < Y < ۱۱۰) \sim ۸۴\ \%}$$

۳) الف) $k>0$

$$E[X] = \sum_{x=1}^{\infty} x\, P(X=x)$$

$$\geq \sum_{x=1}^{k} x\, P(X=x) \geq \sum_{x=1}^{k} x\, P(X=k)$$

چون $x<k \Leftarrow$ … $P(X=x) > P(X=k)$ چون P نزولی است

$$= P(X=k) \sum_{x=1}^{k} x$$

$$= P(X=k) \times \frac{k(k+1)}{2} \geq P(X=k)\,\frac{k^2}{2}$$

$$\Rightarrow E[X] \geq \frac{k^2}{2}\, P(X=k)$$

$$\Rightarrow P(X=k) \leq \frac{2\,E[X]}{k^2}$$

ب) $k \geq 0$

$$E_{[X]} = \int_0^{\infty} x f_{X}(x)\, dx$$

$$\geq \int_0^{k} x f_{X}(x)\, dx \geq \int_0^{k} x f_{X}(k)\, dx$$

چون $x \leq k$ $\Longleftarrow$ $f_{(x)} \geq f_{(k)}$ چون f تابع نزولی است

$$= f_{X}(k) \int_0^{k} x\, dx = f_{X}(k) \frac{k^2}{2}$$

$$\Rightarrow E_{[X]} \geq \frac{k^2}{2} \times f_{X}(k)$$

$$\Rightarrow f_{X}(k) \leq \frac{2 E_{[X]}}{k^2}$$

$$\Rightarrow \boxed{f_{(x)} \leq \frac{2 E_{[X]}}{x^2}}$$

۴) خطا ها را X_i فرض کنید و X می‌شود جمع همه X_i ها که همان تفاوت مجموع حاصل و مجموع اصلی است.

$$X = X_1 + X_2 + X_3 + \cdots + X_{50}$$

$$X_i \sim uniform\left(-\frac{1}{2}, \frac{1}{2}\right)$$

$$P(|X| > 3) = ?$$

$$E[X_i] = 0$$

$$Var[X_i] = \frac{1}{12}$$

طبق قضیه حد مرکزی:

$$Z_n = \frac{\sum_{i=1}^{n} X_i - nE[X_i]}{\sigma\sqrt{n}}$$

$$Z_n \sim normal(0,1)$$

$$Z = \frac{X - 50 \times 0}{\frac{5\sqrt{2}}{2\sqrt{3}}} = \frac{\sqrt{6}}{5} X$$

پس می‌توان تقریب زد که $Z = \frac{\sqrt{6}}{5} X$ یک توزیع نورمال استاندارد است.

$$P(|X| > 3) = P\left(|Z| > \frac{3\sqrt{6}}{5}\right) =$$

$$= 2\Phi\left(-\frac{3\sqrt{6}}{5}\right) = 0/1417$$

پس تقریبا ۱۴ درصد احتمال دارد که این تفاوت بیشتر از ۳ باشد.

۵) فرض کنید آزمایش را n بار تکرار کنیم که از قضیه حد مرکزی داریم :

$$E[X_i] = d = \text{مجهول}$$

$$\operatorname{Var}[X_i] = \sigma^2 = 4$$

$$Z_n = \frac{\sum_{i=1}^{n} X_i - nd}{\sigma\sqrt{n}}$$

که تقریبا Z_n یک توزیع نرمال استاندارد است .

$$P\left(-\frac{1}{2} < \frac{\sum_{i=1}^{n} X_i}{n} - d < \frac{1}{2}\right) =$$

$$P\left(-\frac{1}{2}\frac{\sqrt{n}}{\sigma} < \frac{\sum_{i=1}^{n} X_i - nd}{\sigma\sqrt{n}} < \frac{1}{2}\frac{\sqrt{n}}{\sigma}\right) =$$

$$= P\left(-\frac{\sqrt{n}}{4} < Z_n < \frac{\sqrt{n}}{4}\right) \simeq \Phi\left(\frac{\sqrt{n}}{4}\right) - \Phi\left(-\frac{\sqrt{n}}{4}\right)$$

$$= 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1$$

هرچه n بیشتر باشد تخمین دقیق‌تری زدیم

$$\boxed{P\left(-\frac{1}{2} < \frac{\sum_{i=1}^{n} X_i}{n} - d < \frac{1}{2}\right) \simeq 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1}$$

$$P\left(-\frac{۱}{۲} < \frac{\sum_{i=1}^{n} x_i}{n} - d < \frac{۱}{۲}\right) \geqslant \frac{۹۵}{۱۰۰}$$

$$\rightarrow \quad ۲\phi\left(\frac{\sqrt{n}}{۴}\right) - ۱ \geqslant \frac{۹۵}{۱۰۰}$$

$$\rightarrow \quad \phi\left(\frac{\sqrt{n}}{۴}\right) \geqslant \frac{۳۹}{۴۰}$$

$$\Rightarrow \quad \frac{\sqrt{n}}{۴} \geqslant \frac{۱۹۶}{۱۰۰}$$

$$\Rightarrow \quad n \geqslant ۶۱/۴۷$$

$$\text{و}\ n \in \mathbb{N} \Rightarrow \boxed{n \geqslant ۶۲}$$

یعنی باید حداقل ۶۲ آزمایش انجام دهیم تا با اطمینان ۹۵٪ فاصله‌ی میانگین این آزمایش‌ها از میانگین واقعی کمتر از نصف شود.

ی) برای هر $b>0$ داریم:

$$P(X\geq a) = P(X+b \geq a+b) \leq P\left((X+b)^2 \geq (a+b)^2\right)$$

طبق نامساوی مارکف: $\longleftarrow$ $P(X\geq a) \leq \frac{E[X]}{a}$

$a>0$
X با مقادیر
منفی نیست

$$P(X\geq a) \leq P\left((X+b)^2 \geq (a+b)^2\right) \leq \frac{E[(X+b)^2]}{(a+b)^2}$$

$$= \frac{\sigma^2+b^2}{(a+b)^2}$$

چون برای هر $b>0$ معادله صحیح است فرض کن $b=\frac{\sigma^2}{a}$

$$\rightarrow P(X\geq a) \leq \frac{\sigma^2+\frac{\sigma^4}{a^2}}{\left(a+\frac{\sigma^2}{a}\right)^2} = \frac{a^2\sigma^2+\sigma^4}{(a^2+\sigma^2)^2} = \frac{(a^2+\sigma^2)\sigma^2}{(a^2+\sigma^2)^2}$$

$$= \frac{\sigma^2}{a^2+\sigma^2} \longrightarrow \boxed{P(X\geq a) \leq \frac{\sigma^2}{a^2+\sigma^2} = \frac{Var(X)}{a^2+Var(X)}}$$