



# استنباط آماری

پاییز ۱۴۰۴

## تمرین اول

مرتضی ملکی‌نژاد شوستری - ۸۱۰۱۰۴۲۵۶

### سؤال ۱

$$Cov(X, Y) = E((X - \mu_x)(Y - \mu_y)) = \sum_{i=1}^n \sum_{j=1}^n (x_i - \mu_x)(y_j - \mu_y) P(x_i, y_j) = E(XY) - E(X)E(Y)$$

پس برای محاسبه کوواریانس نیاز است تا  $E(X)$ ,  $E(Y)$ ,  $E(XY)$  حساب شوند.

$$\text{ واضح است که } E(X) = \frac{n+1}{2}$$

برای محاسبه  $E(Y)$  با استفاده از قانون احتمال کل و امید شرطی داریم:

$$E(Y) = \sum_{i=1}^n E(Y|X = x_i)P(X = x_i) = \frac{1}{n} \sum_{i=1}^n E(Y|X = x_i)$$

$$P(Y = y|X = x) = \begin{cases} 0 & (x = y) \\ \frac{1}{n-1} & (x \neq y) \end{cases}$$

$$E(Y|X = x) = \sum_{i=1}^n y_i P(Y = y_i, X = x) = (\sum_{i=1}^n y_i * \frac{1}{n-1}) - \frac{x}{n-1} = \frac{1}{n-1} * (\frac{n(n+1)}{2} - x) \Rightarrow$$

$$E(Y) = \frac{1}{n} \sum_{i=1}^n \frac{1}{n-1} * (\frac{n(n+1)}{2} - x) = \frac{1}{n(n-1)} \sum_{i=1}^n \frac{n(n+1)}{2} - x =$$

$$\frac{1}{n(n-1)} (\frac{n^2(n+1)}{2} - \frac{n(n+1)}{2}) = \frac{1}{n(n-1)} (\frac{n^3 - n}{2}) = \frac{n+1}{2}$$

که مطابق انتظار است زیرا توزیع  $X$  متقارن است و هر تاثیری که از  $x$  به از  $x'$  بزرگ خنثی می‌شود و امید  $Y$  برابر میانگین می‌شود. حال برای محاسبه کوواریانس به محاسبه  $E(XY)$  نیاز است:

$$P(X = x, Y = y) = \begin{cases} 0 & (x = y) \\ \frac{1}{n(n-1)} & (x \neq y) \end{cases}$$

$$E(XY) = \sum_{i=1}^n \sum_{j=1}^n i * j * P(X = i, Y = j) = \frac{1}{n(n-1)} (\sum_{i=1}^n \sum_{j=1}^n ij - \sum_{i=1}^n i^2) =$$

$$\frac{1}{n(n-1)} (\sum_{i=1}^n i)^2 - \sum_{i=1}^n i^2 = \frac{1}{n(n-1)} ((\frac{n(n+1)}{2})^2 - (\frac{n(n+1)(2n+1)}{6})) =$$

$$\frac{1}{n-1} * \left( \frac{n(n+1)^2}{4} - \frac{(n+1)(2n+1)}{6} \right) = \frac{n+1}{n-1} \left( \frac{n^2+n}{4} - \frac{2n+1}{6} \right) = \frac{n+1}{n-1} * \frac{3n^2+3n-4n-2}{12} =$$

$$\frac{n+1}{n-1} * \frac{3n^2-n-2}{12} = \frac{n+1}{n-1} * \frac{(n-1)(3n+2)}{12} = \frac{(n+1)(3n+2)}{12} = \frac{3n^2+5n+2}{12}$$

پس داریم:

$$Cov(X, Y) = E(XY) - E(X)E(Y) = \frac{3n^2+5n+2}{12} - \left(\frac{n+1}{2}\right)^2 = \frac{3n^2+5n+2}{12} - \frac{n^2+2n+1}{4} =$$

$$\frac{-n-1}{12} = -\frac{n+1}{12}$$

## سوال ۲

**a** بخش

$$Y = 5X + \epsilon \Rightarrow \begin{cases} E(Y) = 5E(X) + E(\epsilon) = 5*0 + 0 = 0 \\ Var(Y) = 25Var(X) + Var(\epsilon) = 25 * \frac{4}{12} + 1 = \frac{28}{3} \end{cases}$$

**b** بخش

$$Var(Y) = E(Y^2) - E(Y)^2 \Rightarrow E(Y^2) = Var(Y) + E(Y)^2 = \frac{28}{3} + 0^2 = \frac{28}{3}$$

**c** بخش

$$E(Y|X=x) = E(5X + \epsilon|X=x) = E(5X|X=x) + E(\epsilon|X=x) = 5x + E(\epsilon) = 5x + 0 = 5x$$

**d** بخش

$$E(X^k\epsilon) = E(X^k)E(\epsilon) = 0$$

$$E(X\epsilon^k) = E(X)E(\epsilon^k) = 0$$

$$E(Y^3) = E((5X + \epsilon)^3) \xrightarrow{E(X\epsilon^k)=E(X^k\epsilon)=0} E(Y^3) = 125XE^{(3)} + E(\epsilon^3)$$

از آنجایی که  $x^3$  یک تابع فرد است و توزیع های  $X$  و  $\epsilon$  حول میانگین (صفر) متقارن هستند و بازه انتگرال گیری نیز حول صفر متقارن است می توان نتیجه گرفت که:

$$E(X^3) = E(\epsilon^3) = 0 \Rightarrow E(Y^3) = 0$$

**e** بخش

$$Cov(\epsilon, \epsilon^2) = E(\epsilon^3) - E(\epsilon)E(\epsilon^2) \xrightarrow{E(\epsilon)=0} Cov(\epsilon, \epsilon^2) = E(\epsilon^3)$$

مشابه بخش قبل چون  $\epsilon$  یک توزیع متقارن حول صفر است، امید ممان های فرد آن صفر است پس:

$$E(\epsilon^3) = 0 \Rightarrow Cov(\epsilon, \epsilon^3) = 0$$

ولی این به این معنی نیست که آنها مستقل از هم هستند. به وضوح با دانستن مقدار  $\epsilon$  می توان مقدار  $\epsilon^2$  را بدست آورد. (برای استقلال، دانستن مقدار  $\epsilon$  نباید هیچ دانشی درباره مقدار  $\epsilon^2$  بدهد).

## f بخش

$$Cov(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$$

حال باید دید اگر  $H = aX + b$  باشد، رابطه بین  $X - E(X)$  و  $H - E(H)$  چیست:

$$E(H) = aE(X) + b \Rightarrow H - E(H) = aX + b - aE(X) - b = a(X - E(X))$$

$$\begin{aligned} Cov(H, Y) &= E((H - E(H))(Y - E(Y))) = E(a(X - E(X))(Y - E(Y))) = \\ &aE((X - E(X))(Y - E(Y))) = aCov(X, Y) \end{aligned}$$

با دوبار اعمال این فرمول می‌توان یافت:

$$Cov(n + mX, h + gY) = mCov(X, h + gY) = mgCov(X, Y)$$

## g بخش

$$Cov(X, Y) = E(X - E(X), Y - E(Y))$$

اگر  $M$  را برابر  $X - E(X)$  و  $N$  را برابر  $Y - E(Y)$  قرار داد داریم:

$$E(M) = E(X) - E(E(X)) = 0 \Rightarrow E(M^2) = E(M)^2 + Var(M) = Var(M) = Var(X)$$

$$E(N) = E(Y) - E(E(Y)) = 0 \Rightarrow E(N^2) = E(N)^2 + Var(N) = Var(N) = Var(Y)$$

$$Cov(X, Y) = E(MN)$$

طبق نامساوی کوشی-شوارتز داریم:

$$Cov(X, Y) = E(MN) \leq \sqrt{E(M^2)E(N^2)} = \sqrt{Var(X)Var(Y)} \Rightarrow Cov(X, Y) \leq \sqrt{Var(X)Var(Y)}$$

## سوال ۳

داریم:

$$E(X) = \int_0^\infty xf(x)d_x \xrightarrow{x=\int_0^x 1dt} E(X) = \int_0^\infty \int_0^x dt f(x)d_x$$

چون  $0 < X$  می‌توان ترتیب انتگرال را عوض کرد ولی چون در حدود انتگرال  $x$  وجود دارد اول باید صفحه را برحسب  $t$  بدست آورد:

$$0 \leq x \leq \infty, 0 \leq t \leq x \Rightarrow 0 \leq t \leq \infty, t \leq x \leq \infty$$

پس:

$$\begin{aligned} E(X) &= \int_0^\infty \int_0^x dt f(x)d_x = \int_0^\infty \int_t^\infty f(x)d_x dt \\ \int_t^\infty f(x)d_x &= \int_0^\infty f(x)d_x - \int_0^t f(x)d_x = 1 - F(t) \\ \Rightarrow E(X) &= \int_0^\infty (1 - F(t))dt \end{aligned}$$

## سوال ۴

در جدول زیر، عنوان سطر ها برابر رنگ مشاهده شده توسط شاهد و عنوان ستون ها برابر رنگ واقعی ماشین ها است. هر خانه جدول عنوان می کند که اگر شاهد ماشینی با رنگ برابر با رنگ ستون را ببینید، با چه احتمالی رنگ سطر را می گوید.

جدول ۱: جدول نتایج طبقه بندی

سبز	آبی	
۲	۹۹	آبی
۹۸	۱	سبز

با وجود شهادت شاهد مبنی بر آبی بودن تاکسی، به احتمال  $\frac{2}{101}$  ماشین سبز بوده است. شاهد از هر ۱۰۰ ماشین سبز، ۲ تا را آبی می بیند و احتمال دارد که در شب حادثه نیز یکی از این ماشین های سبز را به اشتباه آبی دیده باشد.

## سوال ۵

$$P(i) = \frac{\lambda^i e^{-\lambda}}{i!} \Rightarrow \frac{P(i)}{P(i-1)} = \frac{\lambda^i e^{-\lambda}}{i!} * \frac{(i-1)!}{\lambda^{i-1} e^{-\lambda}} = \frac{\lambda}{i}$$

با توجه به این، می توان نتیجه گرفت اگر  $\lambda$  بزرگ تر از  $i$  باشد، مقدار  $P(i)$  بزرگ تر از  $P(i-1)$  است. از طرفی واضح است که  $\frac{\lambda}{i}$  با افزایش  $i$  کاهش می یابد. پس باید آخرین جایی را پیدا کرد که این مقدار از یک بزرگ تر است (میدانیم که  $i$  مقداری صحیح و نامنفی دارد):

$$\frac{\lambda}{i} \geq 1 \xrightarrow{i>0} \lambda \geq i \xrightarrow{i \in \mathbb{W}} i \leq \lfloor \lambda \rfloor$$

پس می توان دید آخرین جایی که  $P(i)$  در آن از  $P(i-1)$  بزرگ تر بوده در  $\lfloor \lambda \rfloor = i$  است. به ازای  $i = 0$  نامساوی بالا مبهم می شود. ولی چون هدف یافتن نسبت  $P(i-1)$  به  $P(i)$  است، عملاً جایگذاری  $i = 0$  معنای ندارد و از  $i = 1$  می توان نسبت جملات به همدیگر را یافت.)

## سوال ۶

$$X = \exp(\lambda), Y = \exp(\mu) \Rightarrow f_{X,Y} = f_x f_y = \frac{1}{\lambda \mu} e^{-\left(\frac{x}{\lambda} + \frac{y}{\mu}\right)}$$

$$Z = \min(X, Y)$$

هدف یافتن  $E(Z)$  با توجه به  $X < Y$  است. با توجه به امید شرطی داریم:

$$E(Z|X < Y) = E(\min(X, Y)|X < Y) = E(X|X < Y) = \frac{E(X \cdot 1_{\{X < Y\}})}{P(X < Y)}$$

$$E(X \cdot 1_{\{X < Y\}}) = \int_0^\infty \int_x^\infty x f_{X,Y} dy dx = \int_0^\infty (x * \frac{1}{\lambda} e^{-\frac{x}{\lambda}}) \int_x^\infty \frac{1}{\mu} e^{-\frac{y}{\mu}} dy dx = \int_0^\infty (x * \frac{1}{\lambda} e^{-\frac{x}{\lambda}}) * e^{-\frac{x}{\mu}} \xrightarrow[t=\frac{-x}{\lambda+\mu}]{} =$$

$$\frac{1}{\lambda} \int_0^\infty x e^{-tx} dx = \frac{1}{\lambda} \left[ \frac{-1}{t} x e^{-tx} \right]_0^\infty - \int_0^\infty \frac{1}{t} e^{-tx} dx \xrightarrow{\lim_{x \rightarrow \infty} x e^{-x} = 0} = \frac{1}{\lambda} * \frac{1}{t^2} = \frac{\mu^2 \lambda}{(\lambda + \mu)^2}$$

<sup>۱</sup> این قضیه امید شرطی در درس نبود. با خواندن صفحه ویکی پدیا [۳] و دیدن این ویدیو [۴] فرمول امید شرطی بدست آمد.

$$P(X < Y) = \int_0^\infty \int_x^\infty f_{X,Y} dy dx = \int_0^\infty \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \int_x^\infty \frac{1}{\mu} e^{-\frac{y}{\mu}} dy dx = \int_0^\infty \frac{1}{\lambda} e^{-\frac{x}{\lambda}} * (e^{-\frac{x}{\mu}}) = \frac{1}{\lambda} \int_0^\infty e^{-x(\frac{1}{\lambda} + \frac{1}{\mu})} =$$

$$\frac{1}{\lambda} * \frac{\mu \lambda}{\mu + \lambda} = \frac{\mu}{\mu + \lambda}$$

$$\Rightarrow E(X|X < Y) = \frac{\frac{\lambda \mu^2}{(\lambda + \mu)^2}}{\frac{\mu}{\mu + \lambda}} = \frac{\lambda \mu}{\lambda + \mu}$$

## سوال ۷

با استفاده از قانون امیدهای متوالی می‌توان حساب کرد: (چون به ازای  $X = x$  محاسبات انجام می‌شود، به استقلال متغیرها نیازی نیست.)

$$E((Y - g(X))^2) = E(E([Y - g(x)]^2|X))$$

$$c(x) = E(Y|X = x) \Rightarrow E((Y - g(x))^2|X = x) = g(x)^2 - 2g(x)c(x) + c(x)^2$$

واضح است که مقدار بالا وقتی کمینه است که  $g(x)$  برابر با  $c(x)$  باشد.

## سوال ۸

$$0 \leq y \leq \sqrt{1 - x^2}, -\sqrt{1 - y^2} \leq x \leq \sqrt{1 - y^2}$$

باتوجه به اینکه توزیع توأم  $X$  و  $Y$  یک توزیع یکنواخت است، توزیع  $Y|X = x$  نقطه‌هایی است که در نیم دایره قرار دارند. پس یعنی عملاً نقطه‌های روی پاره خط بین نقاط  $(x, 0)$  و  $(x, \sqrt{1 - x^2})$  هستند. برای پیش‌بینی  $Y$  می‌دانیم بهترین تخمین  $E(Y|X)$  است. باتوجه به این توزیع یکنواخت است و حدود پایین و بالای آن معلوم است داریم:

$$E(Y|X = x) = \frac{0 + \sqrt{1 - x^2}}{2}$$

برای تخمین  $X$  از روی  $Y$  نیز می‌توان به همین طریق حساب کرد با این تفاوت که حد پایین  $X$  صفر نیست.

$$E(X|Y = y) = \frac{-\sqrt{1 - y^2} + \sqrt{1 - y^2}}{2} = 0$$

## سوال ۹

۱

اگر به ازای هر بازه بین  $a_i$  و  $a_{i+1}$  مقدار  $MSE$  را کمینه کرد. می‌توان  $E((X - Y)^2|a_i < X < a_{i+1})$  را حساب کرد و سپس همه این احتمالات را جمع کرد. از طرفی از آنجا که حدود انتگرال‌های  $X$  از هم مجزا می‌شود و  $y_i$  ها تاثیری برهم دیگر ندارند، می‌توان صرفاً بر اساس یکی از این احتمالات شرطی  $MSE$  را کمینه کرد و تعمیم داد. داریم:

$$MSE = E((X - Y)^2) = \sum_{i=-\infty}^{\infty} \int_{a_i}^{a_{i+1}} (x - y_i)^2 f_X(x) f_Y(y) dx \Rightarrow besty_i = argmin \int_{a_i}^{a_{i+1}} (x - y_i)^2 f_X(x) dx$$

از آنجایی که هدف کمینه کردن این مقدار است، می‌توان  $f_Y(y)$  را حذف کرد زیرا مقداری ثابت دارد و سپس از  $y_i$  مشتق گرفت چون نسبت به  $y_i$  مشتق گرفته می‌شود می‌توان مشتق را به درون انتگرال برد. از طرفی چون تابع  $(x - y_i)^2$  نسبت به  $y_i$  یک اکسترمم دارد که مینیمم آن است می‌توان با صفر قرار دادن مشتق مقدار بهینه  $y_i$  را بدست آورد.

$$\begin{aligned} \text{best } y_i &= \underset{\text{argmin}}{\int_{a_i}^{a_{i+1}}} (x - y_i)^2 f_X(x) dx \Rightarrow \int_{a_i}^{a_{i+1}} (2y_i - 2x) f_X(x) dx = 0 \Rightarrow \int_{a_i}^{a_{i+1}} (y_i f_X(x) - x f_X(x)) dx = \\ 0 &\Rightarrow y_i \int_{a_i}^{a_{i+1}} f_X(x) dx = \int_{a_i}^{a_{i+1}} x f_X(x) dx \Rightarrow y_i P(a_i < X < a_{i+1}) = E(X \cdot 1_{a_i < X < a_{i+1}}) \Rightarrow y_i = \\ \frac{E(X \cdot 1_{a_i < X < a_{i+1}})}{P(a_i < X < a_{i+1})} &= E(X | a_i < X < a_{i+1}) \end{aligned}$$

که این دور از انتظار نیست زیرا داشتیم که برای تخمین یک توزیع با یک عدد بهترین عدد برابر میانگین توزیع است. در اینجا نیز چون توزیع  $X$  را بین  $a_i$  و  $a_{i+1}$  بهنظر می‌خواهیم با یک عدد تخمین بزنیم، بهترین تخمین امید  $X$  در آن ناحیه است.

۲

اگر داشته باشیم:

$$\begin{aligned} y_i &= E(X | a_i < X < a_{i+1}) \Rightarrow E(Y | a_i < Y < a_{i+1}) = y_i = E(X | a_i < X < a_{i+1}) \\ \Rightarrow E(X | a_i < X < a_{i+1}) &= E(Y | a_i < Y < a_{i+1}) \end{aligned}$$

از طرفی براساس قانون احتمال کل و قانون امید شرطی داریم:

$$\begin{aligned} E(X) &= E(X | A) + E(X | \bar{A}) \Rightarrow E(X) = \sum_{-\infty}^{\infty} E(a_i < X < a_{i+1}) = \sum_{-\infty}^{\infty} E(a_i < Y < a_{i+1}) = E(Y) \\ \Rightarrow E(X) &= E(Y) \end{aligned}$$

۳

$$\begin{aligned} Var(Y) &= E(Y^2) - E(Y)^2 = E(Y^2) - E(X)^2 \Rightarrow Var(Y) + E((X - Y)^2) - Var(X) = E(Y^2) - E(X)^2 + \\ E((X - Y)^2) - E(X^2) + E(X)^2 &= E(Y^2) + E((X - Y)^2) - E(X^2) = 2E(Y^2) - 2E(XY) \end{aligned}$$

اگر این رابطه به تکه‌های بین نمونه‌برداری شکسته شود داریم:

$$\begin{aligned} Var(Y) + E((X - Y)^2) - Var(X) &= \sigma_{i=-\infty}^{\infty} 2(E(Y^2 | a_i < Y < a_{i+1}) - E(XY | a_i < Y < a_{i+1})) = \\ 2(y_i^2 - E(y_i X | a_i < Y < a_{i+1})) &= 2y_i(y_i - E(X | a_i < Y < a_{i+1})) = 2y_i(y_i - y_i) = 0 \Rightarrow Var(Y) + E((X - Y)^2) - Var(X) = 0 \Rightarrow Var(X) = E((X - Y)^2) + Var(Y) \end{aligned}$$

سوال ۱۰

a

$$\begin{aligned} E(|X - a|) &= \int_0^A |x - a| f_X(x) dx = \frac{1}{A} \int_0^A |x - a| dx = \frac{1}{A} (\int_0^a |x - a| dx + \int_a^A |x - a| dx) = \\ \frac{1}{A} (\int_0^a (a - x) dx + \int_a^A (x - a) dx) &= \frac{1}{A} ([ax - \frac{x^2}{2}]_0^a + [\frac{x^2}{2} - ax]_a^A) = \frac{1}{A} (\frac{a^2}{2} + \frac{A^2}{2} - aA + \frac{a^2}{2}) \end{aligned}$$

هدف کمینه کردن این مقدار است پس داریم:

$$a = \operatorname{argmin} \left( \frac{1}{A} \left( \frac{a^2}{2} + \frac{A^2}{2} - aA + \frac{a^2}{2} \right) \right)$$

$$\frac{d}{da} (2a^2 + A^2 - 2aA) = 0 \Rightarrow 4a - 2A = 0 \Rightarrow a = \frac{A}{2}$$

**b**

مشابه قسمت قبل می‌توان انتگرال را به دو قسمت شکست. البته چون توزیع دیگر یکنواخت نیست نمی‌توان  $f_X(x)$  را از انتگرال بیرون آورد.

$$E(|X - a|) = \int_0^a (a - x) \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx + \int_a^\infty (x - a) \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx$$

چون هدف کمینه کردن این مقدار با تنظیم  $a$  می‌باشد، می‌توان از ضرایب ثابت صرف نظر کرد (در اینجا چون مینیمم بودن اکسترم بدهی نیست باید چک شود):

$$\begin{aligned} \text{best } a &= \operatorname{argmin} \int_0^a (a - x) \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx + \int_a^\infty (x - a) \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx = \operatorname{argmin} \int_0^a (a - x) e^{-\frac{x}{\lambda}} dx + \int_a^\infty (x - a) e^{-\frac{x}{\lambda}} dx \\ \int_0^a (a - x) e^{-\frac{x}{\lambda}} dx &= [-a\lambda e^{-\frac{x}{\lambda}} - -\lambda^2 ((-\frac{1}{\lambda}x - 1)e^{-\frac{x}{\lambda}})]_0^a \xrightarrow[t=\frac{-x}{\lambda}]{} \int_0^a (a - x) e^{-\frac{x}{\lambda}} dx = \\ &-a\lambda e^t + \lambda^2((t - 1)e^t) - (-a\lambda - \lambda^2) \\ \int_a^\infty (x - a) e^{-\frac{x}{\lambda}} dx &= [a\lambda e^{-\frac{x}{\lambda}} - \lambda^2 ((-\frac{1}{\lambda}x - 1)e^{-\frac{x}{\lambda}})]_a^\infty \xrightarrow[t=\frac{-x}{\lambda}]{} = -a\lambda e^t + \lambda^2((t - 1)e^t) \\ \int_0^a (a - x) \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx + \int_a^\infty (x - a) \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx &= -2a\lambda e^t + 2\lambda^2((t - 1)e^t) + a\lambda + \lambda^2 \end{aligned}$$

حال می‌توان نسبت به  $a$  مشتق گرفت و برابر صفر قرار داد (برای راحتی قبل از مشتق همه عوامل بر  $\lambda$  تقسیم شده‌اند).

$$\begin{aligned} -2ae^t + 2\lambda((t - 1)e^t) + a + \lambda &\Rightarrow -2e^t + \frac{-1}{\lambda} * -2a + 2\lambda(\frac{-1}{\lambda})(te^t) + 1 = 0 \Rightarrow \frac{2a}{\lambda} + 1 = 2e^t + 2te^t \Rightarrow \\ 2e^t + 2te^t + 2t &= 1 \end{aligned}$$

این معادله با روش‌های کلی حل نمی‌شود و با استفاده از روش‌های عددی باید حل شود. با استفاده از کتابخانه‌های محاسبه مقدار  $t$  حدوداً برابر

$$t = \frac{-x}{\lambda} - 0.181955 \text{ می‌شود و از آنجا که } a = 0.81955\lambda \text{ برابر می‌شود.}$$

**سوال ۱۱**

**a** بخش

- $X \sim Uniform(0, 2)$

$$X \sim Uniform(0, 2) \Rightarrow \mu = 1, \sigma^2 = \frac{(2 - 0)^2}{12} = \frac{1}{3} \Rightarrow \sigma = \frac{\sqrt{3}}{3}$$

$$P(|X - \mu| \geq 2\sigma) = P(|X - 1| \geq \frac{2\sqrt{3}}{3}) = P(X > 1, X - 1 > \frac{2\sqrt{3}}{3}) + P(X < 1, 1 - X < \frac{-2\sqrt{3}}{3})$$

$$= 2P(X > 1 + \frac{\sqrt{3}}{3}) \xrightarrow{1+2-\frac{3}{3} \approx 2.15} 2P(X > 1 + \frac{2\sqrt{3}}{3}) = 2 * 0 = 0$$

که با نامساوی چبیشف نیز همخوانی دارد.

- $P(-1) = P(1) = \frac{1}{2}$

$$\mu = 0, \sigma^2 = E((X - \mu)^2) = 1 \Rightarrow \sigma = 1$$

$$P(|X - \mu| > 2\sigma) = P(|X| > 2) = P(X > 2) + P(X < -2) = 0$$

## b بخش

$$P(|X - \mu| > c) \leq \frac{E((X - \mu)^4)}{c^4}$$

$$\textbf{Markov} \Rightarrow P((X - \mu)^4 > c) < \frac{E((X - \mu)^4)}{c^4}$$

از طرفی داریم:

$$|X - \mu| > c \xrightarrow{|X - \mu| > c > 0} (|X - \mu|)^2 > c^2 \Rightarrow (X - \mu)^4 > c^4$$

یعنی اگر  $x - \mu$  از  $c$  بزرگتر باشد،  $(x - \mu)^4$  هم از  $c^4$  بزرگتر است. پس می‌توان نتیجه گرفت:

$$P(|X - \mu| > c) < P((X - \mu)^4 > c^4)$$

اگر این دونامساوی را کنار هم گذاشت داریم:

$$\begin{cases} P((X - \mu)^4 > c) < \frac{E((X - \mu)^4)}{c^4} \\ P(|X - \mu| > c) < P((X - \mu)^4 > c^4) \end{cases} \Rightarrow P(|X - \mu| > c) < \frac{E((X - \mu)^4)}{c^4}$$

## c بخش

حل نشده.

## سوال ۱۲

درواقع هدف محاسبه این مقدار است:

$$E(\exp(En(X)))$$

به صورت کلی برای توزیع نمایی با پارامتر  $\lambda$  داریم:

$$f_{\exp(\lambda)}(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \Rightarrow E(\exp(\lambda)) = \lambda$$

پس برای توزیع  $(En(X))$  داریم:

$$E(f(g(x))) = \int f(g(x))f_X(x)dx = \int e^{en(x)} \frac{1}{\sqrt{2\pi}} e^{-x^2} = \frac{1}{\sqrt{2\pi}} \int e^{en(x)-x^2}$$

(این سوال با این فرض حل شد که تابع توزیع چگالی احتمال توزیع نمایی مطابق چیزی که در اسلایدها گفته شده یعنی  $\frac{1}{\lambda}e^{-\frac{x}{\lambda}}$  باشد.)

## سوال ۱۳

بخش‌های خواسته شده در کد اکسریپت آمده‌اند. ولی برای شفاف‌سازی در گزارش مسئله به صورت تحلیلی بررسی می‌شود. جهت سادگی می‌توان فرض کرد که مقدار شرط‌بندی در هر قسمت مقدار ثابت «یک» می‌باشد. زیرا حالات دیگر را می‌توان به این حالت تبدیل کرد. صرفاً کافی است که تبدیل‌های زیر انجام شوند:

$$\text{target amount} = \lceil \frac{\text{target amount} - (\text{initial amount} \bmod \text{bet amount})}{\text{bet amount}} \rceil$$

$$\text{initial amount} = \lfloor \frac{\text{initial amount}}{\text{bet amount}} \rfloor$$

$$\text{bet amount} = 1$$

پس مسئله به این تغییر می‌کند که با چه احتمالی تعداد برد‌های قمارباز از تعداد باخت‌های او به اندازه  $\text{target amount} - \text{initial amount}$  بیشتر است به شرطی که تعداد باخت‌های او از ابتدای بازی بیشتر از تعداد برد‌ها به علاوه  $\text{initial amount}$  نشود. می‌توان بالاستفاده از توزیع دوچمله‌ای فرم باز احتمال برد را نوشت. ولی از آنجا که هدف این سوال این نیست و زمان زیادی می‌طلبد با یک جستجو می‌توان فرم بسته راه حل را یافت. [۱]

## سوال ۱۴

فرمول کشیدگی<sup>۲</sup> با جستجو پیدا شد. [۴] به غیر از این تنها مورد اضافه کشیدن نمودارهای ویولنی بر روی نمودارهای جعبه‌ای است که توزیع تاخیرها را بهتر نمایش می‌دهد.

برای الگوریتم‌ها می‌توان دید که با این که مرتب‌سازی ادغامی و مرتب‌سازی سریع به مرتب‌سازی درجی هستند، اما کشیدگی مرتب‌سازی درجی کمتر است. زیرا برای محاسبه کشیدگی داده‌ها با توجه به مقدار میانگین و انحراف معیار نرمال می‌شوند و با توجه به نمودار ویولنی می‌توان دید که اگر مقیاس درنظر گرفته نشود (زیرا نرمال می‌شود) در مرتب‌سازی درجی، داده‌ها فاصله کمتری از میانگین دارند. البته در مرتب‌سازی سریع اگر ایده‌هایی مثل انتخاب دو عضو برای pivot و سه تکه کردن مرتب‌سازی و یا حذف Tail recursion انجام شود، می‌توان به یک الگوریتم پایدارتر<sup>۳</sup> دست یافت.

به چهار الگوریتم معروفی شده، الگوریتم مرتب‌سازی پایتون TimSort [۵] نام دارد که ایده کلی آن این است که آرایه‌های کوچک با استفاده از مرتب‌سازی درجی و آرایه‌های بزرگ با استفاده از مرتب‌سازی ادغامی مرتب شوند تا در  $n$  های کوچک، از ضریب کم مرتب‌سازی ادغامی بتوان استفاده کرد و در  $n$  های بزرگ، بتوان با اردر nlogn همچنان آرایه را مرتب نمود. پس در اینجا نیز انتظار می‌رود که سرعت Timsort شبیه مرتب‌سازی ادغامی و مرتب‌سازی سریع باشد (که درواقع از آنجایی که در هسته و با C پیاده‌سازی شده سریع‌تر نیز است). و شکل توزیع آن شبیه مرتب‌سازی ادغامی باشد.

<sup>2</sup>Kurtosis

<sup>3</sup>Consistent

خواسته سوال اعمال نویز<sup>۴</sup> با توجه به توزیع داده شده است. طبق توضیحات داده شده، توزیع باید بر روی مکان پیکسل نویز تاثیر بزارد و نه بر روی تغییر رنگ پیکسل. به همین منظور برای توزیع‌های نرمال و نمایی مقدار  $X$  و  $Y$  برابر دو نمونه‌برداری یک بعدی از توزیع مادر هستند که در کنار هم قرار می‌گیرند و پیکسل‌های نویزی را تشکیل می‌دهند. همچنین از آنجایی که توزیع نمایی به شدت کاهشی است، برای آن که نویز نمایش معناداری روی تصویر پیدا نکند، با توجه به میزان کمینه و بیشینه در نمونه‌برداری، نقطه‌های نمونه‌برداری شده به مقیاس تصویر رسیدند تا بتوان بهتر شکل توزیع را نشان داد (در غیر اینصورت صرفاً در پنج پیکسل ابتدایی تصویر جمعیت پیکسل‌ها را می‌داد که داده‌ای نمی‌دهد).

### بخش a

برای این بخش مطابق خواسته نویز بر روی تصاویر اعمال شد. برای توزیع یکنواخت پیکسل‌های نویزی تمرکز خاصی روی یک قسمت تصویر ندارند و در همه تصویر پراکنده هستند.

برای توزیع نمایی نیز تمرکز پیکسل نویز بر بخش بالا سمت چپ تصویر (مبدا مختصات) است، زیرا در توزیع نمایی با فاصله گرفتن از مبدا احتمال چگالی نقطه بصورت نمایی کم می‌شود.

در بخش توزیع نرمال، از آنجایی که میانگین و واریانس بصورت تصادفی و یکنواخت تعیین می‌شود نمی‌توان حالت کلی بیان کرد. اگر واریانس قسمت  $X$  و قسمت  $Y$  باهم برابر می‌بود، نویز به شکل یک دایره محو شونده به وجود می‌آمد که میزان شدت محوشدنگی به میزان واریانس و جایگاه مرکز دایره به مقدار میانگین  $X$  و  $Y$  بستگی می‌داشت. برای واریانس بالا نویز نرمال به نویز یکنواخت نزدیک می‌شود و برای واریانس پایین اکثر تصویر سالم است و صرفاً یک بخش کوچک از تصویر به شدت درگیر نویز می‌شود. در مثال‌های نمایش داده شده می‌توان این موضوع را برای انتخاب‌های مختلف پارامترهای توزیع دید.

### بخش b

خیر میزان اطلاعات از دست رفته به مکانی که نویز در آن اعمال می‌شود بستگی دارد. درواقع خواسته سوال محاسبه احتمال زیر است:

$$E(a < X < b, c < Y < d)$$

برای مثال برای توزیع نمایی که تمرکز بیشتری در مبدا دارد، تقریباً هیچ داده‌ای از دست نمی‌رود، زیرا فاصله صورت از مبدا به نسبت زیاد است (با توجه به روند شدید کاهشی توزیع نمایی) با توجه بهتابع CDF داریم:

$$E(a < X < b, c < Y < d) = [1 - e^{\frac{1}{x}}]^b_a * [1 - e^{\frac{1}{x}}]^d_c = (e^{\frac{1}{a}} - e^{\frac{1}{b}}) * (e^{\frac{1}{c}} - e^{\frac{1}{d}})$$

با تغییر کمینه و بیشینه از حالت معمول به ابتدا و انتهای تصویر، صرفاً مقدار  $a, b, c, d$ ، تغییر می‌کند و شکل توزیع تغییر نمی‌کند. اگر ضرب مقیاس  $k$  باشد، همه این مقادیر در  $k$  ضرب خواهند شد و سپس باید مقایسه کرد که حاصل ضرب این احتمال در میزان نویز درنهایت بالای 0.2 می‌شود یا خیر (که با شهود می‌شود دید که خیر و از آنجا که سوال مقدار دقیق خواسته حساب نمی‌شود).

برای توزیع نرمال، اگر پارامترها باعث بشوند که تمرکز نویز روی صورت نباشد، میزان اطلاعات از دست رفته می‌تواند تغییر کند. چون تابع CDF توزیع نرمال فرم بسته‌ای ندارد، و سوال نیز نخواسته، فرمول احتمال محاسبه نمی‌شود اما شهودا می‌توان دریافت که احتمال از دست رفته اطلاعات از توزیع نمایی بیشتر است ولی اگر میانگین توزیع از قسمت صورت دور باشد و یا واریانس زیاد به طوری که تمرکز پیکسل‌های خراب روی یک قسمت نباشد و یا واریانس کم به طوری که پیکسل‌های خراب وارد صورت نشوند احتمال از دست رفتن اطلاعات کمتر است.

برای توزیع یکنواخت از آنجا که پیکسل‌های خراب در هیچ قسمتی تمرکز ندارند و بصورت یکنواخت پخش شده‌اند، به راحتی می‌توان دریافت که نسبت پیکسل‌های خراب به همه پیکسل‌ها برابر همان نرخ نویز می‌باشد. به طور خاص برای حالت مرزی ۲۰٪ چون ممکن است در نمونه برداری

<sup>4</sup>noise

پیکسل‌های مرزی به صورت کاملاً یکنواخت نباشند، احتمال خرابی تصویر به سمت ۰.۵ سوق پیدا می‌کند ولی برای باقی موارد این احتمال مطابق انتظار برابر ۱ می‌باشد.

## بخش ۶

به طور کلی فیلتر میانه‌گیر برای حذف نویز نمک و فلفل عملکرد بهتری نسبت به فیلترهایی مثل فیلتر میانگین‌گیری و یا فیلتر گاووسی دارد. زیرا تاثیر نویزهای خراب را کمینه می‌کند. پارامتر مهم فیلتر میانه‌گیر سایز کرنل می‌باشد. با زیاد شدن سایز کرنل، نویزهای نمک و فلفل بیشتری حذف می‌شوند ولی اطلاعات رنگی پیکسل‌ها نیز بیشتر از بین می‌روند و با کرنل‌های بزرگ، تعداد رنگ‌های تصویر کم شده و تصویر حالت آبرنگی پیدا می‌کند.

فیلتر میانه گیر در حالتی که پیکسل‌های نویز نمک و فلفل گستردگی بالایی داشته باشند بهترین عملکرد خود را دارد زیرا با تجمع پیکسل‌های نویزی، ممکن است میانه آن‌ها نیز روی یکی از نقاط نویز بیفتند ولی با گستردگی بالا، می‌توان با احتمال خوبی مطمئن بود که میانه نویز نیست و یکی از نقاط اصلی تصویر است.

## منابع خارج از درس

- [1] Naman Agrawal. “Gambler’s Ruin Problem: A Probability Antiquity”. In: *Medium* (2022). Accessed: 2025-11-20.
- [2] Bright side of maths. *Probability Theory 21 | Conditional Expectation given events*. <https://www.youtube.com/watch?v=GDQW5AC0SDQ>. Accessed: 2025-11-20.
- [3] Wikipedia contributors. *Conditional Expectation*. [https://en.wikipedia.org/wiki/Conditional\\_expectation](https://en.wikipedia.org/wiki/Conditional_expectation). Accessed: 2025-11-20.
- [4] Wikipedia contributors. *Kurtosis*. <https://en.wikipedia.org/wiki/Kurtosis>. Accessed: 2025-11-20.
- [5] Wikipedia contributors. *Timsort*. <https://en.wikipedia.org/wiki/Timsort>. Accessed: 2025-11-20.