

امید ریاضی تابعی از متغیرهای تصادفی و کوواریانس

فردوس گرگی

امید ریاضی

➤ امید ریاضی متغیر تصادفی گسسته:

تعریف: اگر متغیر تصادفی گسسته X دارای تابع جرم احتمال $f(x)$ باشد، میانگین یا مقدار مورد انتظار یا امید ریاضی X برابر است با:

$$E(X) = \sum_x x f(x)$$

➤ امید ریاضی متغیر تصادفی پیوسته:

تعریف: اگر متغیر تصادفی پیوسته X دارای تابع چگالی احتمال $f(x)$ باشد، میانگین یا مقدار مورد انتظار یا امید ریاضی X برابر است با:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

➤ قضیه:

فرض کنید X یک متغیر تصادفی گسسته با توزیع احتمال $f(x)$ باشد. میانگین یا مقدار مورد انتظار متغیر تصادفی $g(X)$ عبارت است از:

$$E(g(X)) = \sum_x g(x) f(x)$$

فرض کنید X یک متغیر تصادفی پیوسته با چگالی احتمال $f(x)$ باشد. میانگین یا مقدار مورد انتظار متغیر تصادفی $g(X)$ عبارت است از:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$$

➤ امید ریاضی :

فرض کنید X و Y دو متغیر تصادفی با تابع توزیع احتمال توأم $f(x,y)$ باشند.
میانگین یا مقدار مورد انتظار متغیر تصادفی $g(X,Y)$ برابر است با:

$$\mu_{g(X,Y)} = E(g(X,Y)) = \sum_x \sum_y g(x,y) f(x,y) \quad \text{در حالت گسسته:}$$

$$\mu_{g(X,Y)} = E(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) dx dy \quad \text{در حالت پیوسته:}$$

➤ مثال ۱: تابع توزیع احتمال توأم مقابل را در نظر بگیرید.
مطلوب است $E(X - Y)$ و $E(XY)$

X \ Y	1	2	3
1	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{9}$
2	0	$\frac{1}{6}$	$\frac{1}{9}$
3	0	0	$\frac{1}{9}$

$$E(XY) = \sum_x \sum_y xyf(x, y) = (1)(1)\left(\frac{1}{3}\right) + (2)(1)\left(\frac{1}{6}\right) + (3)(1)\left(\frac{1}{9}\right) + (1)(2)(0) + (2)(2)\left(\frac{1}{6}\right) + (3)(2)\left(\frac{1}{9}\right) + (1)(3)(0) + (2)(3)(0) + (3)(3)\left(\frac{1}{9}\right) = \frac{10}{3}$$

$$E(X - Y) = \sum_x \sum_y (x - y)f(x, y) = (1 - 1)\left(\frac{1}{3}\right) + (2 - 1)\left(\frac{1}{6}\right) + (3 - 1)\left(\frac{1}{9}\right) + (1 - 2)(0) + (2 - 2)\left(\frac{1}{6}\right) + (3 - 2)\left(\frac{1}{9}\right) + (1 - 3)(0) + (2 - 3)(0) + (3 - 3)\left(\frac{1}{9}\right) = \frac{4}{9}$$

➤ **مثال ۲:** فرض کنید شرکتی بسته‌های یک کیلویی شامل دو نوع شکلات و یک نوع آبنبات را بسته‌بندی می‌کند و لزوماً وزن سه نوع محصول در جعبه‌ها یکسان نیست. اگر X را وزن شکلات نوع اول و Y را وزن شکلات نوع دوم در نظر بگیریم، تابع چگالی احتمال توأم آن به صورت زیر

$$f(x,y) = \begin{cases} 24xy & 0 < x < 1; \ 0 < y < 1 \\ & x + y \leq 1 \\ 0 & \text{o.w} \end{cases} \quad \text{است:}$$

مقدار مورد انتظار وزن مجموع شکلات‌ها چقدر است؟

$$\begin{aligned} E(X + Y) &= \int_0^1 \int_0^{1-y} (x + y) \times 24xy \, dx dy \\ &= \int_0^1 \int_0^{1-y} (24x^2y + 24xy^2) \, dx dy = \int_0^1 \left[(8x^3y + 12x^2y^2) \Big|_0^{1-y} \right] dy \\ &= \int_0^1 (4y^4 - 12y^2 + 8y) dy = \left(\frac{4}{5}y^5 - 4y^3 + 4y^2 \right) \Big|_0^1 = \frac{4}{5} \end{aligned}$$

➤ **مثال ۳:** تابع چگالی احتمال توأم زیر را در نظر بگیرید. اگر $z = \sqrt{x^2 + y^2}$ ، مطلوب است محاسبه $E(z)$.

$$f(x, y) = \begin{cases} 4xy & 0 < x < 1 ; 0 < y < 1 \\ 0 & o.w \end{cases}$$

$$\begin{aligned} E(Z) &= E\left(\sqrt{x^2 + y^2}\right) = \int_0^1 \int_0^1 (\sqrt{x^2 + y^2}) \times 4xy \, dx dy \\ &= \int_0^1 2y \int_0^1 2x \sqrt{x^2 + y^2} \, dx dy = \int_0^1 2y \left[\frac{2}{3} (x^2 + y^2)^{\frac{3}{2}} \right]_0^1 dy \\ &= \frac{2}{3} \int_0^1 (2y(1 + y^2)^{\frac{3}{2}} - 2y^4) dy = \frac{2}{3} \left(\frac{2}{5} (1 + y^2)^{\frac{5}{2}} - \frac{2}{5} y^5 \right) \Big|_0^1 = \frac{8}{15} (2\sqrt{2} - 1) \\ &= \mathbf{0.9752} \end{aligned}$$

➤ **قضیه:** مقدار مورد انتظار مجموع یا تفاضل دو یا چند تابع از متغیرهای تصادفی X و Y برابر با مجموع یا تفاضل مقادیر مورد انتظار توابع مذکور است. یعنی:

$$E[g(X, Y) \pm h(X, Y)] = E(g(X, Y)) \pm E(h(X, Y))$$

اثبات: (در حالت پیوسته)

$$\begin{aligned} E[g(X, Y) \pm h(X, Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (g(x, y) \pm h(x, y)) f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy \pm \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy \\ &= E(g(X, Y)) \pm E(h(X, Y)) \end{aligned}$$

➤ **نتیجه ۱:** اگر $h(X,Y) = h(Y)$ و $g(X,Y) = g(X)$ ، آنگاه :

$$E(g(X) \pm h(Y)) = E(g(X)) \pm E(h(Y))$$

➤ **نتیجه ۲:** اگر $h(X,Y) = Y$ و $g(X,Y) = X$ ، آنگاه :

$$E(X \pm Y) = E(X) \pm E(Y)$$

➤ **نکته:** اگر $g(X,Y) = X$ ، داریم :

$$E(g(X,Y)) = E(X)$$

$$E(X) = \begin{cases} \sum_x \sum_y xf(x,y) = \sum_x x \overbrace{\sum_y f(x,y)}^{f_X(x)} = \sum_x xf(X) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x,y) dy dx = \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f(x,y) dy}_{f_X(x)} dx = \int_{-\infty}^{\infty} xf(x) dx \end{cases}$$

➤ کوواریانس

اگر قرار دهیم $g(X,Y) = (X - \mu_X)(Y - \mu_Y)$ ، مقدار مورد انتظار $g(X,Y)$ کمیتی به نام کوواریانس را بدست می‌دهد که میزان وابستگی X و Y به یکدیگر را نشان می‌دهد.

تعریف : فرض کنید X و Y متغیرهای تصادفی با تابع توزیع احتمال توأم $f(x,y)$ باشد. کوواریانس X و Y عبارت است از:

حالت گسسته:

$$\sigma_{XY} = cov(x, y) = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y)$$

حالت پیوسته:

$$\sigma_{XY} = cov(x, y) = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy$$

* میزان وابستگی بین X و Y را نشان می‌دهد.

* ممکن است مثبت، منفی یا صفر باشد.

* اگر X و Y مستقل باشند، کوواریانس آن‌ها صفر است ولی عکس این موضوع لزوماً برقرار نیست.

$$\sigma_{XY} = E(XY) - E(X)E(Y) = E(XY) - \mu_X\mu_Y$$

اثبات: (در حالت پیوسته)

$$\begin{aligned}\sigma_{XY} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy \\&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy - \mu_X y - \mu_Y x + \mu_X \mu_Y) f(x, y) dx dy \\&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \underbrace{xy f(x, y) dx dy}_{E(XY)} - \mu_X \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f(x, y) dx}_{f_Y(y)} dy \\&\quad - \mu_Y \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f(x, y) dy}_{f_X(x)} dx + \mu_X \mu_Y \underbrace{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy}_1 = E(XY) - \mu_X \mu_Y\end{aligned}$$

Y \ X	1	2	3	$f_Y(y)$
1	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{11}{18}$
2	0	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{5}{18}$
3	0	0	$\frac{1}{9}$	$\frac{1}{9}$
$f_X(x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

➤ **مثال ۴:** تابع توزیع احتمال توأم مقابل را در نظر بگیرید.
مطلوب است کوواریانس X و Y .

در مثال ۱ محاسبه کردیم که $E(XY) = \frac{10}{3}$
داریم:

$$\mu_X = E(X) = \sum_x x f(x) = (1)\left(\frac{1}{3}\right) + (2)\left(\frac{1}{3}\right) + (3)\left(\frac{1}{3}\right) = 2$$

$$\mu_Y = E(Y) = \sum_y y f(y) = (1)\left(\frac{11}{18}\right) + (2)\left(\frac{5}{18}\right) + (3)\left(\frac{1}{9}\right) = \frac{3}{2}$$

$$\sigma_{XY} = E(XY) - E(X)E(Y) = \frac{10}{3} - (2)\left(\frac{3}{2}\right) = \frac{1}{3} > 0 \Rightarrow$$

اگر X زیاد شود، انتظار می‌رود Y نیز زیاد شود. اگر Y زیاد شود، انتظار داریم X هم زیاد شود

➤ **مثال ۵:** تابع چگالی احتمال زیر را در نظر بگیرید. مطلوبست محاسبه‌ی کوواریانس X و Y .

$$f(x, y) = \begin{cases} 24xy & 0 < x < 1; 0 < y < 1 \\ & x + y \leq 1 \\ 0 & o.w \end{cases}$$

$$\begin{aligned} E(XY) &= \int_0^1 \int_0^{1-x} 24x^2y^2 dydx = \int_0^1 8x^2(1-x^3)dx \\ &= \left(\frac{8x^3}{3} - 6x^4 + \frac{24x^5}{5} - \frac{8x^6}{6} \right) \bigg|_0^1 = \frac{4}{30} \end{aligned}$$

$$\begin{aligned} E(X) &= \int_0^1 \int_0^{1-x} 24x^2y dydx = \int_0^1 12x^2(1-x)^2 dx \\ &= \left(4x^3 - 6x^4 + \frac{12x^5}{5} \right) \bigg|_0^1 = \frac{2}{5} \end{aligned}$$

$$E(Y) = \frac{2}{5}$$

$$\sigma_{XY} = E(XY) - E(X)E(Y) = \frac{4}{30} - \left(\frac{2}{5}\right)\left(\frac{2}{5}\right) = \frac{-2}{75} < 0$$

اگر شکلات نوع اول زیاد شود، انتظار داریم شکلات نوع دوم کم شود و بالعکس.

➤ ضریب همبستگی

تعریف: فرض کنید X و Y متغیرهای تصادفی با کوواریانس σ_{XY} و انحراف معیارهای σ_X و σ_Y باشند. ضریب همبستگی X و Y عبارت است از:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

* واحد ندارد

* $-1 \leq \rho_{XY} \leq 1$ (نزدیک صفر \equiv همبستگی کم ؛ نزدیک ۱ و -۱ \equiv همبستگی زیاد)

* مقیاس اعداد در آن اثر ندارد

* اگر Y ترکیب خطی از X باشد ($Y = aX + b$) ، آنگاه:

$$\rho_{XY} = \begin{cases} 1 & a > 0 \\ -1 & a < 0 \end{cases}$$

➤ **مثال ۶:** در مثال ۴، ضریب همبستگی را بیابید.

از قبل محاسبه کردیم $\mu_Y = \frac{3}{2}$ ، $\mu_X = 2$ ، $\sigma_{XY} = \frac{1}{3}$ داریم:

$$E(X^2) = \sum_x x^2 f(x) = (1)^2 \left(\frac{1}{3}\right) + (2)^2 \left(\frac{1}{3}\right) + (3)^2 \left(\frac{1}{3}\right) = \frac{14}{3}$$

$$\sigma_X^2 = E(X^2) - (E(X))^2 = \frac{14}{3} - 2^2 = \frac{2}{3} \quad \Rightarrow \quad \sigma_X = \frac{\sqrt{2}}{\sqrt{3}}$$

$$E(Y^2) = \sum_y y^2 f(y) = (1)^2 \left(\frac{11}{18}\right) + (2)^2 \left(\frac{15}{18}\right) + (3)^2 \left(\frac{2}{18}\right) = \frac{49}{18}$$

$$\sigma_Y^2 = E(Y^2) - (E(Y))^2 = \frac{49}{18} - \left(\frac{3}{2}\right)^2 = \frac{17}{36} \quad \Rightarrow \quad \sigma_Y = \frac{\sqrt{17}}{6}$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\frac{1}{3}}{\frac{\sqrt{2}}{\sqrt{3}} \times \frac{\sqrt{17}}{6}} = \frac{2\sqrt{3}}{\sqrt{2}\sqrt{17}} = \sqrt{\frac{12}{34}}$$

➤ **مثال ۷:** در مثال ۵، ضریب همبستگی را محاسبه کنید.

از قبل محاسبه کردیم $\sigma_{XY} = \frac{-2}{75}$ ، $E(X) = \frac{2}{5}$ ، $E(Y) = \frac{2}{5}$. داریم:

$$E(X^2) = \int_0^1 \int_0^{1-x} 24x^3 y \, dy dx = \int_0^1 [12x^3 y^2 \Big|_0^{1-x}] dx = \int_0^1 12x^3 (1-x)^2 dx \\ = \left(3x^4 - \frac{24x^5}{5} + 2x^6 \right) \Big|_0^1 = \frac{1}{5}$$

$$\sigma_X^2 = E(X^2) - (E(X))^2 = \frac{1}{5} - \left(\frac{2}{5}\right)^2 = \frac{1}{25} \quad \Rightarrow \quad \sigma_X = \frac{1}{5}$$

به طریق مشابه داریم $\sigma_Y = \frac{1}{5}$. لذا:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\frac{-2}{75}}{\frac{1}{5} \times \frac{1}{5}} = -\frac{2}{3}$$

➤ **قضیه:** فرض کنید X و Y دو متغیر تصادفی مستقل باشند. در این صورت:

$$E(XY) = E(X)E(Y) \quad \text{یا} \quad \sigma_{XY} = 0$$

اثبات: (در حالت پیوسته)

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) \, dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_X(x)f_Y(y) \, dx dy \\ &= \int_{-\infty}^{\infty} yf_Y(y) \int_{-\infty}^{\infty} xf_X(x) \, dx dy = \mu_X \int_{-\infty}^{\infty} yf_Y(y) \, dy = \mu_X \mu_Y = E(X)E(Y) \end{aligned}$$

نتیجه: اگر X و Y مستقل باشند، $\rho_{XY} = 0$.

➤ **قضیه:** اگر X و Y متغیرهای تصادفی با تابع توزیع احتمال توأم $f(x,y)$ باشند، آنگاه:

$$\sigma^2_{aX+bY} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{XY}$$

اثبات:

$$\begin{aligned}\sigma^2_{aX+bY} &= E((aX + bY)^2) - (E(aX + bY))^2 = E(a^2X^2 + 2abXY + b^2Y^2) \\ &\quad - (aE(x) + bE(y))^2 = \underline{a^2E(X^2)} + \underline{2abE(XY)} + \underline{b^2E(Y^2)} - \underline{a^2(E(X))^2} \\ &\quad - \underline{2abE(X)E(Y)} - \underline{b^2(E(Y))^2} = a^2 \left(E(X^2) - (E(X))^2 \right) + b^2 (E(Y^2) - (E(Y))^2) \\ &\quad + 2ab(E(XY) - E(X)E(Y)) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY}\end{aligned}$$

➤ **نکته:** اگر X و Y متغیرهای تصادفی با تابع توزیع احتمال توأم $f(x,y)$ باشند، آنگاه:

$$\sigma^2_{aX+bY+c} = \sigma^2_{Z+c} = \sigma^2_Z$$

\uparrow
 $Z = aX + bY$ قرار می‌دهیم

نتیجه: اگر X و Y مستقل باشند، آنگاه:

$$\sigma^2_{aX \pm bY} = a^2 \sigma^2_X + b^2 \sigma^2_Y$$

و اگر X_n, \dots, X_1 مستقل باشند، آنگاه:

$$\sigma^2_{a_1X_1+\dots+a_nX_n} = a_1^2 \sigma^2_{X_1} + a_2^2 \sigma^2_{X_2} + \dots + a_n^2 \sigma^2_{X_n}$$

➤ نکته :

$$* \sigma_{(\sum_{i=1}^n a_i X_i)}^2 = \sigma_{a_1 X_1 + \dots + a_n X_n}^2 = \sum_{i=1}^n a_i^2 \sigma_{X_i}^2 + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} a_i a_j \sigma_{X_i X_j}$$

$$* Cov(X, a) = 0 \quad E(ax) - E(a)E(X) = aE(X) - aE(X) = 0$$

$$* Cov(X, X) = Var(X) \quad E(XX) - E(X)E(X) = E(X^2) - (E(X))^2 = \sigma_X^2$$

$$* Cov(X, Y) = Cov(Y, X) \quad \sigma_{XY} = E(XY) - E(X)E(Y) = E(YX) - E(Y)E(X) = \sigma_{YX}$$

$$* Cov(aX + bY, cZ + dW) = ac\sigma_{XZ} + ad\sigma_{XW} + bc\sigma_{YZ} + bd\sigma_{YW}$$

$$* Cov(X + t, Z + s) = Cov(X, Z)$$



مثال ۸: فرض کنید شرکتی دو بازاریاب دارد که مستقل از یکدیگر و به صورت پورسانتی حقوق دریافت می‌کنند. بازاریاب اول به طور متوسط، ماهانه ۱۰۰ دلار (با انحراف معیار ۱۰ دلار) و بازاریاب دوم ماهانه به طور متوسط ۱۵۰ دلار (با انحراف معیار ۲۰ دلار) دریافت می‌کند. اگر شرکت به اندازه ۱۰٪ حقوق هر بازاریاب به شرکت بیمه پرداخت کند، میانگین و انحراف معیار مبلغی که شرکت بابت بازاریاب‌ها هزینه می‌کند چقدر است؟

$$\begin{aligned} E(X) &= 100 & \sigma_X &= 10 \\ E(Y) &= 150 & \sigma_Y &= 20 \end{aligned}$$

میزان هزینه شرکت برای بازاریاب‌ها
↓

$$Z_1 = 1.10 X \quad ; \quad Z_2 = 1.10 Y \quad \quad Z = Z_1 + Z_2 = 1.1 X + 1.1 Y$$

$$E(Z) = 1.1E(X) + 1.1E(Y) = 275 \quad \text{دلار} \quad \begin{matrix} X \text{ و } Y \text{ مستقل} \\ \nearrow 0 \end{matrix}$$

$$\begin{aligned} \sigma_Z^2 &= \sigma_{1.1X+1.1Y}^2 = 1.21\sigma_X^2 + 1.21\sigma_Y^2 + (1.1)(1.1)\cancel{\sigma_{XY}} \\ &= 1.21 \times 100 + 1.21 \times 400 = 605 \quad \Rightarrow \quad \sigma_Z = \sqrt{605} = 24.6 \quad \text{دلار} \end{aligned}$$

➤ **مثال ۹:** فرض کنید X_1, X_2, X_3 دارای واریانس‌های ۵ و ۴ و ۷ باشند و نیز $\sigma_{X_1 X_2} = 3$ ، $\sigma_{X_1 X_3} = -1$ و X_2, X_3 مستقل باشند. اگر $Y_1 = X_1 - 2X_2 + 3X_3$ و $Y_2 = -2X_1 + 3X_2 + 4X_3$ مطلوب $Var(Y_1)$ و $Cov(Y_1, Y_2)$ است.

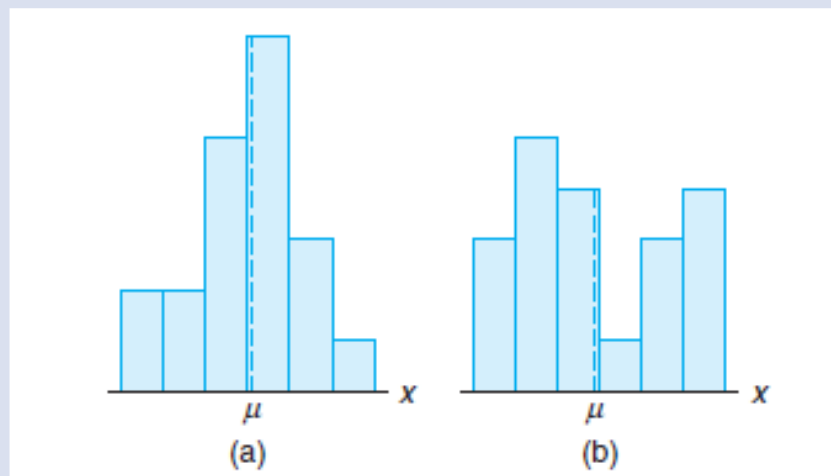
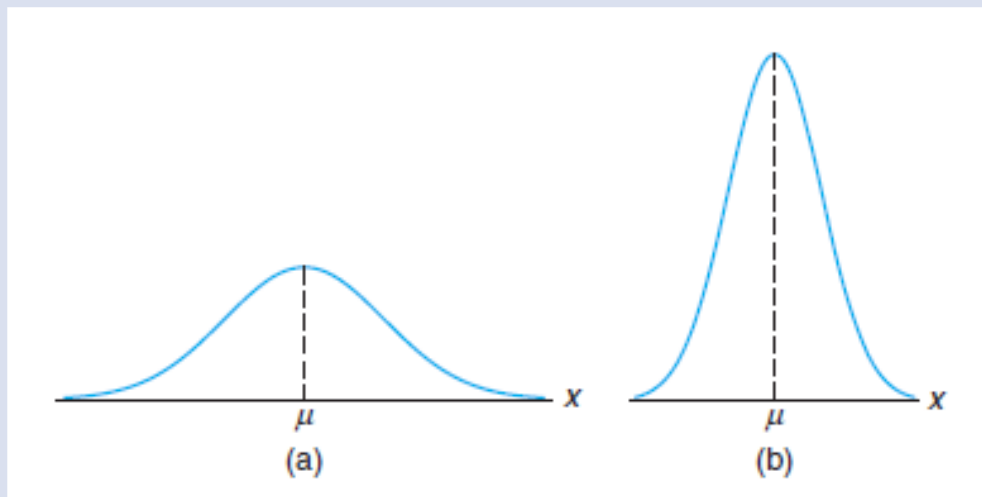
$$\begin{aligned} Var(Y_1) &= Var(X_1 - 2X_2 + 3X_3) = \sigma_{X_1}^2 + 4\sigma_{X_2}^2 + 9\sigma_{X_3}^2 - 4\sigma_{X_1 X_2} + 6\sigma_{X_1 X_3} - 12\sigma_{X_2 X_3} \\ &= 5 + 4(4) + 9(7) - 4(3) + 6(-1) + 12(0) = \mathbf{66} \end{aligned}$$

$$\begin{aligned} Cov(Y_1, Y_2) &= Cov((-2X_1 + 3X_2) + 4X_3, (X_1 - 2X_2) + 3X_3) \\ &= Cov(-2X_1 + 3X_2, X_1 - 2X_2) + 3Cov(-2X_1 + 3X_2, X_3) + 4Cov(X_3, X_1 - 2X_2) \\ &\quad + 12Cov(X_3, X_3) = -2Cov(X_1, X_1) + 4Cov(X_1, X_2) + 3Cov(X_2, X_1) - 6Cov(X_2, X_2) \\ &\quad - 6Cov(X_1, X_3) + 9Cov(X_2, X_3) + 4Cov(X_3, X_1) - 8Cov(X_3, X_2) \\ &\quad + 12\sigma_{X_3}^2 = -2\sigma_{X_1}^2 + 7\sigma_{X_1 X_2} - 6\sigma_{X_2}^2 - 2\sigma_{X_1 X_3} + \sigma_{X_2 X_3} + 12\sigma_{X_3}^2 \\ &= -2(5) + 7(3) - 6(4) - 2(-1) + 1(0) + 12(7) = \mathbf{73} \end{aligned}$$

➤ **قضیه چبیشف:** احتمال این که متغیر تصادفی دلخواه X مقداری بین k برابر انحراف معیار از میانگین را اختیار کند، حداقل $1 - \frac{1}{k^2}$ است، یعنی

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

در واقع احتمال این که مشاهدات در فاصله متقارنی حول میانگین توزیع باشد، به مقدار انحراف معیار توزیع (میزان پراکندگی داده‌ها) وابسته است. مثلاً می‌دانیم که همواره حداقل $\frac{3}{4}$ داده‌ها در هر توزیعی در فاصله کمتر از 2σ از میانگین توزیع هستند. این قضیه برای هر توزیع دلخواهی از مشاهدات برقرار است و تنها یک کران پایین برای احتمال به دست می‌دهد نه مقدار واقعی احتمال را. فقط زمانی که توزیع معلوم باشد می‌توان مقدار دقیق احتمال مورد نظر را به دست آورد. لذا این قضیه تنها در حالت‌هایی که شکل توزیع نامعلوم است و فقط میانگین و انحراف معیار را می‌دانیم به کار می‌رود.



➤ **مثال ۱۰:** فرض کنید متغیر تصادفی X دارای میانگین $\mu = 12$ و واریانس $\sigma^2 = 9$ بوده و توزیع احتمال آن نامعلوم است. مقادیر زیر را بیابید:

$$P(6 < X < 18), \quad P(|X - 12| \geq 9).$$

$$P(6 < X < 18) = P(12 - (3)(2) < X < 12 + (3)(2)) \geq 1 - \frac{1}{4} = \frac{3}{4}$$

$$\begin{aligned} P(|X - 12| \geq 9) &= 1 - P(|X - 12| < 9) \\ &= 1 - P(12 - (3)(3) < X < 12 + (3)(3)) \leq 1 - (1 - \frac{1}{9}) = \frac{1}{9} \end{aligned}$$