חוקי דה-מורגן :

$$\left(\bigcap_{k=1}^{n} A_k\right)^C = \bigcup_{k=1}^{n} A_k^C \qquad\qquad \left(\bigcup_{k=1}^{n} A_k\right)^C = \bigcap_{k=1}^{n} A_k^C$$

<u>הגדרה – הסתברות מותנית</u>

$$\mathbb{P}\{A \mid B\} \triangleq \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}$$

נוסחת בייס ($Bayes$) :

$$\mathbb{P}\{A \mid B\} = \mathbb{P}\{B \mid A\}\frac{\mathbb{P}\{A\}}{\mathbb{P}\{B\}}$$

נוסחת ההסתברות הכוללת :

$$\bigcup_{n=0}^{\infty} B_n = \Omega \quad \Rightarrow \quad \mathbb{P}\{A\} = \sum_{n=0}^{\infty} \mathbb{P}\{A \mid B_n\}\mathbb{P}\{B_n\}$$

<u>הגדרה – חוסר תלות סטטיסטית</u>

מאורעות יקראו בת"ס (בלתי-תלויים סטטיסטית) כאשר :

$$\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\}\mathbb{P}\{B\}$$

משתנים אקראיים בדידים מוכרים :

כינוי ; הסבר	סימון	פונקצית הסתברות	תוחלת (μ)	שונות (σ^2)	הערות
ברנולי ; הצלחה בניסוי	$Bern(p)$	$\begin{cases} p, & 1 \\ q, & 0 \end{cases}$	p	pq	$q \triangleq 1-p$
בינומי ; k הצלחות מתוך n ניסויים	$Bin(n,p)$	$\binom{n}{k} p^k q^{n-k}$ $k = 0,1,2,\ldots,n$	np	npq	כאשר n גדול, ו p קטן אזי ניתן לקרב $Bin(n,p) \approx Pois(np)$
גיאומטרית ; k כשלונות עד הצלחה ראשונה בניסויים	$Geom(p)$	pq^{k-1} $k = 1,2,3,\ldots$	$\frac{1}{p}$	$\frac{q}{p^2}$	
פואסונית	$Pois(\lambda)$	$e^{-\lambda}\frac{\lambda^k}{k!}$ $k = 0,1,2,\ldots$	λ	λ	$Pois(\lambda_1) + Pois(\lambda_2)$ $= Pois(\lambda_1 + \lambda_2)$
מולטי-נומית	$Mult(n,\vec{p})$	$\frac{n!}{x_1!\cdots x_n!}\prod_{i=1}^{n} p_i^{x_i}$	$\mu_{X_i} = np_i$		

משתנים אקראיים רציפים מוכרים:

כינוי	סימון	פונקצית צפיפות	תוחלת (μ)	שונות (σ^2)	הערות
מעריכית	$Exp(\lambda)$	$\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$ $\lambda>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$m_k=\frac{k!}{\lambda^k}$
גאוסיאנית (נורמאלית) סטנדרטית	$N(0,1)$	$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$	0	1	$m_{2k}=(2k-1)!!$ $m_{2k+1}=0$
גאוסיאנית (נורמאלית) כללית	$N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2	עבור $X\sim N(\mu,\sigma^2)$ ו $Z\sim N(0,1)$ מתקיים: $P(X\le b)=P\left(Z\le\frac{b-\mu}{\sigma}\right)$
אחידה	$U[a,b]$	$\begin{cases}\frac{1}{b-a}, & x\in[a,b]\\ 0, & x\notin[a,b]\end{cases}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	
גאמה	$Gamma(r,\lambda)$ $\Gamma(r,\lambda)$	$\frac{\lambda^r}{(r-1)!}x^{r-1}e^{-\lambda x}$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	r פעמים קונבולוציה של $Exp(\lambda)$ עם עצמו

משתנים אקראיים

פונקצית התפלגות:

$$F_X(x)\triangleq P(X\le x),\quad f_X(x)=\frac{d}{dx}F_X(x)$$

צפיפות של סכום משתנים אקראיים היא קונבולוציה בין הצפיפויות שלהם:

$$f_{X=U+V}(x)=f_U(u)*f_V(v)\triangleq\int_{-\infty}^{\infty}f_U(u)f_V(x-u)\,du$$

תכונות:

$$\Gamma(r,\lambda)+\Gamma(s,\lambda)=\Gamma(r+s,\lambda)$$

$$Pois(\lambda_1)+Pois(\lambda_2)=Pois(\lambda_1+\lambda_2)$$

$$Bin(n,p)+Bin(m,p)=Bin(n+m,p)$$

תכונת חוסר הזיכרון:

$$X\sim Geom(p)\quad\Rightarrow\quad P(X>n+k\mid X>n)=P(X>k)$$

$$X\sim Exp(\lambda)\quad\Rightarrow\quad P(X>s+t\mid X>t)=P(X>s)$$

צפיפות של משתנה מקרי שתלוי במשתנה מקרי אחר, כאשר $Y=h(X)$:

$$f_Y(y)=\left.\frac{f_X(x)}{\left|h'(x)\right|}\right|_{x=h^{-1}(y)}$$

התוחלת

הגדרה – תוחלת

תוחלת של משתנה אקראי בדיד

$$E[N] \triangleq \sum_{n=-\infty}^{\infty} n \cdot p_N(n)$$

תוחלת של משתנה אקראי רציף

$$E[X] \triangleq \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx$$

הרחבת המושג: תוחלת של פונקציה של משתנה:

עבור משתנה אקראי בדיד ופונקציה $f(x)$:

$$E[f(N)] \triangleq \sum_{n=-\infty}^{\infty} f(n) \cdot p_N(n)$$

עבור משתנה אקראי רציף ופונקציה $h(x)$:

$$E[h(x)] = \int_{-\infty}^{\infty} h(x) \cdot f_X(x)\,dx$$

תכונות התוחלת ותוצאות נוספות

1. ליניאריות: עבור a,b קבועים דטרמיניסטים:

$$E[aX+bY] = aE[X]+bE[Y]$$

2. נוסחת התוחלת הכוללת:

$$E[A] = E[A\mid B]P(B) + E\left[A\mid B^C\right]P\left(B^C\right)$$

3. משפט ההחלקה:

$$E[X] = E\left[E[X\mid Y]\right]$$

4. אם X,Y בלתי תלויים (או אפילו רק חסרי קורלציה), אז

$$E[XY] = E[X]\cdot E[Y]$$

5. לכל מ"א X רציף (קיים ביטוי מקביל למשתנים בדידים):

$$E[X] = \int_0^{\infty} \mathbb{P}\{X<x\}\,dx = \int_0^{\infty} \left(1-F_X(x)\right)dx$$

הגדרה – מומנט

$$m_k \triangleq E\left[X^k\right] = \int_{-\infty}^{\infty} x^k f_X(x)\,dx$$

מומנטים של $Z \sim N\left(0,\sigma^2\right)$:

$$E[Z^n] = \begin{cases} \sigma^n \cdot 1\cdot 3\cdot \ldots \cdot (n-1), & n=2k \\ 0, & n=2k+1 \end{cases} = \begin{cases} \sigma^n \cdot n!!, & n=2k \\ 0, & n=2k+1 \end{cases}$$

הגדרה – שונות

$$\operatorname{var} X = \sigma_X^2 \triangleq E\left[\left(X-E[X]\right)^2\right] = E\left[X^2\right] - \left(E[X]\right)^2$$

תכונות השונות, כאשר c קבוע דטרמיניסטי (לא אקראי):

$$\operatorname{var}(c) = 0$$

$$\operatorname{var}(cX) = c^2 \operatorname{var} X$$

$$\operatorname{var}(X+Y) = \operatorname{var} X + 2\operatorname{cov}(X,Y) + \operatorname{var} Y = \sigma_X^2 + 2\sigma_{XY} + \sigma_Y^2$$

הגדרה – סטיית התקן

$$\sigma_X \triangleq \sqrt{\operatorname{var} X}$$

הגדרה – פונקציה יוצרת מומנטים (פונקציה אופיינית)

$$M_X(s) \triangleq Ee^{sX} \quad \Rightarrow \quad M_X^{(k)}(0) = EX^k = m_k$$

אי שוויון צ'בישב:

$$\mathbb{P}\{|X-\mu| > b\} \le \frac{\sigma^2}{b^2}, \qquad \mathbb{P}\{|X-\mu| > a\sigma\} \le \frac{1}{a^2}$$

אי שוויון ינסן:

$$E[h(X)] \ge h(E[X])$$

וקטור אקראי

הגדרה – וקטור אקראי

וקטור אקראי ממימד n הוא וקטור אשר כל n רכיביו הם משתנים אקראיים, ומסומן:

$$\vec{X} = (X_1, X_2, \ldots, X_n)$$

מצפיפות משותפת של שני משתנים, ניתן לחשב את פונקצית הצפיפות השולית (של אחד המשתנים):

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

אי תלות קיימת אם"ם

$$f_{XY}(x,y) = f_X(x) f_Y(y)$$

אי תלות בין רכיבי וקטור (X,Y) לא יכולה להתקיים אם התמך של $f_{XY}(x,y)$ לא מלבני.

התפלגות מותנית:

$$f_{X|Y}(x \mid y) \triangleq \frac{f_{XY}(x,y)}{f_Y(y)} \underset{Bayes}{=} f_{Y|X}(y \mid x)\frac{f_X(x)}{f_Y(y)}$$

נוסחת הצפיפות הכוללת:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x \mid y) f_Y(y)\,dy$$

טרנספורמציה של וקטור אקראי

$$\begin{cases} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f(x_1,\ldots,x_n) \\ T(x_1,\ldots,x_n) \xrightarrow{1:1} (y_1,\ldots,y_n) \end{cases} \Rightarrow \begin{cases} y_1 = h_1(x_1,\ldots,x_n) \\ \vdots \\ y_n = h_n(x_1,\ldots,x_n) \end{cases}$$

$$\Rightarrow \quad f_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n) = \left|\frac{\partial \vec{x}}{\partial \vec{y}}\right| f_{X_1,\ldots,X_n}\left(T^{-1}(y_1,\ldots,y_n)\right) = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial y_n} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} f_{X_1,\ldots,X_n}\left(T^{-1}(y_1,\ldots,y_n)\right)$$

שונות בין שני משתנים אקראיים (קו-ווריאנס):

$$\sigma_{XY} = \operatorname{cov}(X,Y) \triangleq E\big[(X - E[X])(Y - E[Y])\big] = E[XY] - E[X]E[Y]$$

תכונות הקווריאנס, עבור a,b קבועים דטרמיניסטים:

$$\operatorname{cov}(X,X) = \operatorname{var} X = \sigma_X^2$$

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

$$\operatorname{cov}(aX+b,Y) = a\operatorname{cov}(X,Y)$$

קורלציה (מקדם המתאם):

$$\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y},\ |\rho_{XY}|\le 1$$

תכונות:

$$\rho_{X,Y}=\rho_{Y,X}$$
$$\rho_{aX,Y}=sign(a)\,\rho_{X,Y}$$
$$\rho_{X+c,Y}=\rho_{X,Y}$$

וקטור אקראי גאוסי

עבור וקטור אקראי גאוסי כללי $\vec{X}$, הצפיפות המשותפת היא

$$f_{\vec{X}}(\vec{x})=C\exp\left\{-\frac{1}{2}\left((\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T\right)\right\}$$

כאשר $\vec{\mu}$ הוא וקטור התוחלות, ו Σ מטריצת הקוווריאנס (תמיד סימטרית).

ועבור מקרה פרטי של שלושה משתנים:

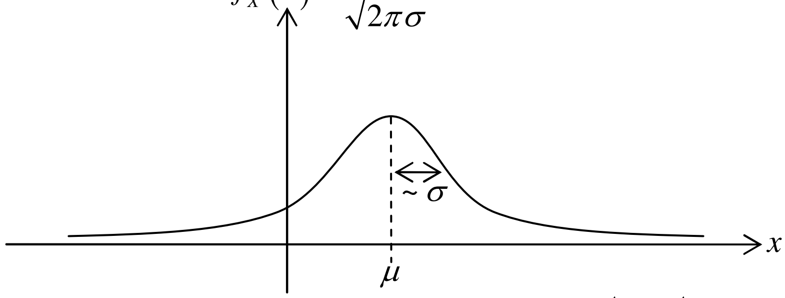
$$f_{X,Y,Z}(x,y,z)=C\cdot\exp\left\{-\frac{1}{2}\begin{pmatrix}x-\mu_X & y-\mu_Y & z-\mu_Z\end{pmatrix}\begin{pmatrix}\sigma_X^2 & \sigma_{XY} & \sigma_{XZ}\\ \sigma_{XY} & \sigma_Y^2 & \sigma_{YZ}\\ \sigma_{XZ} & \sigma_{YZ} & \sigma_Z^2\end{pmatrix}^{-1}\begin{pmatrix}x-\mu_X\\ y-\mu_Y\\ z-\mu_Z\end{pmatrix}\right\}$$

בהינתן k מימד הוקטור $\vec{X}$, הקבוע C הוא

$$C=\frac{1}{\sqrt{(2\pi)^k|\Sigma|}}\exp\left\{-\frac{1}{2}\vec{\mu}\Sigma^{-1}\mu^T\right\}$$

למציאת התוחלת של משתנה גאוסי, יש לגזור את צפיפות המשתנה ולהשוות ל 0, מכיוון שתוחלת משתנה גאוסי מתקבלת במקסימום המקומי שלו:

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



ולכן, כדי לקבל את וקטור התוחלות של וקטור גאוסי, נדרוש

$$\vec{\nabla}f_{X,Y,Z}(x,y,z)=0$$

<u>טענה שימושית</u>

בהינתן וקטור גאוסי עם מטריצת קוווריאנס Σ_{XYZ}, וק״ל של (x,y,z):

$$u=\vec{u}_1\cdot(x,y,z)$$
$$v=\vec{v}_1\cdot(x,y,z)$$

אזי

$$\Sigma_{UV}=\begin{pmatrix}\vec{u}_1\\ \vec{v}_1\end{pmatrix}\Sigma_{XYZ}\begin{pmatrix}\vec{u}_1 & \vec{v}_1\end{pmatrix}$$

במשפחה הגאוסית, כלומר עבור משתנים אקראיים גאוסים המהווים חלק מוקטור אקראי גאוסי, משתנים בלתי מתואמים (חסרי קורלציה) הם גם בלתי תלויים.

חזאים

החזאי (משערך) האופטמלי של Y בעזרת X (פונקציה של X):

$$\hat{Y}_{opt} = E[Y \mid X]$$

החזאי (משערך) הליניארי האופטמלי של Y בעזרת X (פונקציה ליניארית של X):

$$\hat{Y}_{opt}^{lin} = \frac{\operatorname{cov}(X,Y)}{\sigma_X^2}(X - EX) + EY$$

במשפחה הגאוסית, כלומר עבור משתנים אקראיים גאוסים המהווים חלק מוקטור אקראי גאוסי, החזאי האופטמלי הוא החזאי הליניארי האופטמלי.

החוק החלש של המספרים הגדולים ($WLLN - Weak\ Law\ of\ Large\ Numbers$):

יהיו $X_1, X_2, \dots, X_N$ משתנים אקראיים מפולגים באופן זהה ובת"ל ($IID - Independent\ Indentically\ Distributed$), בעלי תוחלת μ. נסמן $S_n = \sum_{i=1}^{n} X_i$, אזי לכל $\varepsilon > 0$:

$$\lim_{n\to\infty} \mathbb{P}\left\{\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right\} = 0$$

בעצם, חוק זה אומר כי אם נגדיר $X = \frac{S_n}{n}$, אזי $\lim_{n\to\infty} f_X(x) = \delta(x-\mu)$

משפט הגבול המרכזי ($CLT - Central\ Limit\ Theorem$):

יהיו $X_1, X_2, \dots, X_N$ משתנים אקראיים מפולגים באופן זהה ובת"ל ($IID - Independent\ Indentically\ Distributed$), בעלי תוחלת μ ושונות σ^2. נסמן $S_n = \sum_{i=1}^{n} X_i$, אזי

$$\lim_{n\to\infty} \mathbb{P}\left\{\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le a\right\} = \Phi(a)$$

כאשר Φ פונקצית ההתפלגות של $N(0,1)$.