פרק 7: תכונות של תוחלת (סיכום) **(20416 / 21.7.10)**

טענות:

1. **תוחלת של פונקציה של X ו- Y.** נניח כי $g(x,y)$ היא פונקציה ממשית של x ו- y.

אם ל-X ו-Y יש פונקציית הסתברות משותפת $p_{X,Y}$, אז: $$E[g(X,Y)]=\sum_y\sum_x g(x,y)\,p_{X,Y}(x,y)$$

אם ל-X ו-Y יש פונקציית צפיפות משותפת $f_{X,Y}$, אז: $$E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,f_{X,Y}(x,y)\,dx\,dy$$

2. **תוחלת של סכום משתנים מקריים**

אם $E[X_i]$ סופית לכל $i=1,2,...,n$ אז: $$E[X_1+X_2+...+X_n]=E[X_1]+E[X_2]+\ ...\ +E[X_n]$$

הערה: הטענה במקרה האינסופי תקפה, אם כל ה-X_i-ים הם משתנים מקריים אי-שליליים, או אם $\sum_{i=1}^{\infty}E[|X_i|]<\infty$.

השונות המשותפת של X ו-Y מוגדרת על-ידי: $$\mathrm{Cov}(X,Y)=E[(X-E[X])(Y-E[Y])]$$ $$=E[XY]-E[X]E[Y]$$

אם $\mathrm{Cov}(X,Y)=0$ אז X ו-Y נקראים **בלתי-מתואמים**.

תכונות השונות המשותפת:
1. $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
2. $\mathrm{Cov}(X,X)=\mathrm{Var}(X)$
3. $\mathrm{Cov}(aX,Y)=a\,\mathrm{Cov}(X,Y)$
4. $\mathrm{Cov}\left(\sum_{i=1}^{n}X_i,\sum_{j=1}^{m}Y_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}\mathrm{Cov}(X_i,Y_j)$

טענות:
1. אם X ו-Y משתנים מקריים בלתי-תלויים, אז לכל שתי פונקציות ממשיות g ו-h מתקיים: $E[g(X)h(Y)]=E[g(X)]E[h(Y)]$
2. אם X ו-Y בלתי-תלויים זה בזה, אז הם בלתי-מתואמים (אך לא בהכרח להיפך).
3. אם $\underline{X}\sim Mult(n,\underline{p})$ אז לכל $i\neq j$ מתקיים $\mathrm{Cov}(X_i,X_j)=-np_ip_j$.

שונות של סכום סופי של משתנים מקריים

$$\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right)=\mathrm{Cov}\left(\sum_{i=1}^{n}X_i\,,\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathrm{Var}(X_i)+2\sum_{i<j}\mathrm{Cov}(X_i,X_j)=\sum_{i=1}^{n}\mathrm{Var}(X_i)+\sum_{i\neq j}\mathrm{Cov}(X_i,X_j)$$

שונות של סכום סופי של משתנים מקריים בלתי-תלויים $$\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathrm{Var}(X_i)$$

מקדם המתאם של X ו-Y מוגדר על-ידי: $$\rho(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

טענות:
1. לכל X ו-Y מתקיים $-1\le\rho(X,Y)\le 1$.
2. אם $Y=aX+b$ אז $\rho(X,Y)=\begin{cases}1\ , & a>0\\ -1\ , & a<0\end{cases}$.

אינדיקטור הוא משתנה מקרי שמקבל את הערכים 1 ו-0 בהסתברויות p $(0<p<1)$ ו- $1-p$, בהתאמה.

לכל אינדיקטור, שנסמנו ב-X, מתקיים: $E[X]=p \quad ; \quad E[X^2]=p \quad ; \quad \mathrm{Var}(X)=p(1-p)$

ולכל שני אינדיקטורים, שנסמנם ב-X_1 וב-X_2, מתקיים: $E[X_1X_2]=P\{X_1=1, X_2=1\}$

ולכן: $\mathrm{Cov}(X_1,X_2)=P\{X_1=1,X_2=1\}-P\{X_1=1\}P\{X_2=1\}$

תוחלת מותנית

אם ל-X ו-Y יש פונקציית הסתברות משותפת אז **התוחלת המותנית של X בהינתן $Y=y$** מוגדרת, לכל y המקיים $P\{Y=y\}>0$, על-ידי:

$$E[X\,|\,Y=y]=\sum_x xP\{X=x\,|\,Y=y\}=\sum_x x\frac{P\{X=x,Y=y\}}{P\{Y=y\}}$$

אם ל-X ו-Y יש פונקציית צפיפות משותפת אז **התוחלת המותנית של X בהינתן $Y=y$** מוגדרת, לכל y המקיים $f_Y(y)>0$, על-ידי:

$$E[X\,|\,Y=y]=\int_{-\infty}^{\infty} x\,f_{X|Y}(x\,|\,y)\,dx=\int_{-\infty}^{\infty} x\,\frac{f_{X,Y}(x,y)}{f_Y(y)}\,dx$$

הערה: התוחלת המותנית של X בהינתן $Y=y$ היא פונקציה של הערך y.

תוחלת של תוחלת מותנית: $E[X]=E[E[X\,|\,Y]]$

כלומר, אם Y משתנה מקרי בדיד אז – $E[X]=\sum_y E[X\,|\,Y=y]P\{Y=y\}$

ואם Y משתנה מקרי רציף אז – $E[X]=\int_{-\infty}^{\infty} E[X\,|\,Y=y]f_Y(y)\,dy$

אפשר להשתמש בטענה זו, כדי לחשב הסתברויות של מאורעות על-ידי התניה במשתנה מקרי.

נסמן מאורע כלשהו ב-E: אם Y משתנה מקרי בדיד אז – $P(E)=\sum_y P(E\,|\,Y=y)P\{Y=y\}$

ואם Y משתנה מקרי רציף אז – $P(E)=\int_{-\infty}^{\infty} P(E\,|\,Y=y)\,f_Y(y)\,dy$

טענה (תרגיל ת26): $E[g(Y)X]=E[g(Y)E[X\,|\,Y]]$

נוסחת השונות המותנית: $\mathrm{Var}(X)=E[\mathrm{Var}(X\,|\,Y)]+\mathrm{Var}(E[X\,|\,Y])$

סכום מקרי

אם N הוא משתנה מקרי בדיד שערכיו שלמים אי-שליליים, ואם X_1, X_2, ... הם משתנים מקריים שווי-התפלגות ובלתי-תלויים זה בזה וב-N, אז:

$$E\left[\sum_{i=1}^{N}X_i\right]=E[N]E[X_1]$$

[כאשר $N=0$, סכום המשתנים שווה גם הוא ל-0]

$$\mathrm{Var}\left(\sum_{i=1}^{N}X_i\right)=E[N]\mathrm{Var}(X_1)+(E[X_1])^2\mathrm{Var}(N)$$

הפונקציה יוצרת המומנטים של משתנה מקרי X מוגדרת לכל t ממשי, על-ידי : $M_X(t) = E[e^{tX}]$

המומנט מסדר r של משתנה מקרי X מוגדר על-ידי : $E[X^r]$

המומנט מסדר r של משתנה מקרי X מתקבל מהנגזרת ה- r-ית (לפי t) של $M_X(t)$ בנקודה $t=0$, אם הנגזרת קיימת בסביבת נקודה זו. כלומר, מתקיים השוויון : $M_X^{(r)}(t)\big|_{t=0} = E[X^r]$

פונקציה יוצרת מומנטים של סכום מקרי של מ"מ ב"ת ש"ה $Y = \sum_{i=1}^{N} X_i$: $M_Y(t) = E\left[\left(M_{X_1}(t)\right)^N\right]$

פונקציות יוצרות מומנטים של התפלגויות מיוחדות

משתנה מקרי בינומי : $X \sim B(n,p)$ $\quad M_X(t) = (pe^t + 1 - p)^n$, לכל t

משתנה מקרי פואסוני : $X \sim Po(\lambda)$ $\quad M_X(t) = e^{\lambda(e^t - 1)}$, לכל t

משתנה מקרי גיאומטרי : $X \sim Geo(p)$ $\quad M_X(t) = \dfrac{pe^t}{1-(1-p)e^t}$, $t < -\ln(1-p)$

משתנה מקרי בינומי שלילי : $X \sim NB(r,p)$ $\quad M_X(t) = \left(\dfrac{pe^t}{1-(1-p)e^t}\right)^r$, $t < -\ln(1-p)$

משתנה מקרי אחיד : $X \sim U(a,b)$ $\quad M_X(t) = \dfrac{e^{bt} - e^{at}}{t(b-a)}$, לכל $t \neq 0$

משתנה מקרי מעריכי : $X \sim Exp(\lambda)$ $\quad M_X(t) = \dfrac{\lambda}{\lambda - t}$, $t < \lambda$

משתנה מקרי גמא : $X \sim Gamma(s,\lambda)$ $\quad M_X(t) = \left(\dfrac{\lambda}{\lambda - t}\right)^s$, $t < \lambda$

משתנה מקרי נורמלי סטנדרטי : $X \sim N(0,1)$ $\quad M_X(t) = e^{t^2/2}$, לכל t

משתנה מקרי נורמלי : $X \sim N(\mu,\sigma^2)$ $\quad M_X(t) = e^{\mu t + \sigma^2 t^2/2}$, לכל t

משתנה מקרי חי-בריבוע : $X \sim Chi\text{-}square(n)$ $\quad M_X(t) = \left(\dfrac{1}{1-2t}\right)^{n/2}$, $t < 1/2$

טענות: **1.** אם $Y = aX + b$, אז : $M_Y(t) = M_{aX+b}(t) = e^{bt} M_X(at)$

2. אם $X_1, X_2, \ldots, X_n$ הם משתנים מקריים בלתי-תלויים, אז : $M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t)$

3. קיימת התאמה חד-חד-ערכית בין הפונקציות יוצרות המומנטים לפונקציות ההתפלגות המצטברת. כלומר, אם הפונקציה יוצרת המומנטים של משתנה מקרי X קיימת וסופית בסביבה כלשהי של $t=0$, אז ההתפלגות של X נקבעת באופן יחיד.

תוצאות: **1.** סכום של n משתנים מקריים בינומיים בלתי-תלויים עם הפרמטרים n_i ו- p, לכל $i = 1,...,n$, הוא משתנה מקרי בינומי עם הפרמטרים Σn_i ו- p .

2. סכום של n משתנים מקריים פואסוניים בלתי-תלויים עם הפרמטרים λ_i, לכל $i = 1,...,n$, הוא משתנה מקרי פואסוני עם הפרמטר $\Sigma\lambda_i$.

3. סכום של n משתנים מקריים גיאומטריים בלתי-תלויים שלכולם הפרמטר p , הוא משתנה מקרי בינומי שלילי עם הפרמטרים n ו- p .

4. סכום של n משתנים מקריים נורמליים בלתי-תלויים עם הפרמטרים μ_i ו- σ_i^2 , הוא משתנה מקרי נורמלי עם הפרמטרים $\Sigma\mu_i$ ו- $\Sigma\sigma_i^2$.

5. סכום של n משתנים מקריים מעריכיים בלתי-תלויים שלכולם הפרמטר λ, הוא משתנה מקרי גמא עם הפרמטרים n ו- λ.

הפונקציה יוצרת המומנטים המשותפת של המשתנים המקריים X_1, X_2, ... , X_n מוגדרת לכל n ערכים ממשיים t_1 , t_2 , ... , t_n, על-ידי :

$$M_{\underline{X}}(t_1,\ldots,t_n) = E\left[e^{t_1 X_1+\ldots+t_n X_n}\right]$$

הפונקציה יוצרת המומנטים המשותפת של משתנים מקריים קובעת באופן יחיד את התפלגותם המשותפת.

ההתפלגות הרב-נורמלית

יהיו Z_1 , Z_2 , ... , Z_n משתנים מקריים נורמליים סטנדרטיים בלתי-תלויים.

אם לקבועים מסוימים a_{ij} ו-μ_i, כאשר $i = 1,...,m$ ו- $j = 1,...,n$, מתקיים :

$$\begin{aligned} X_1 &= a_{11}Z_1 + \ldots + a_{1n}Z_n + \mu_1 \\ X_2 &= a_{21}Z_1 + \ldots + a_{2n}Z_n + \mu_2 \\ &\vdots \\ X_m &= a_{m1}Z_1 + \ldots + a_{mn}Z_n + \mu_m \end{aligned}$$

אז למשתנים המקריים X_1 , X_2 , ... , X_m יש התפלגות **רב-נורמלית**.

טענות: **1.** פונקציית הצפיפות המשותפת של X_1, X_2, ... , X_m מוגדרת לכל $\underline{x}$, שהוא וקטור של m מספרים ממשיים, על-ידי :

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{m/2}\sqrt{|det\,\Sigma|}} \cdot e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^t \Sigma^{-1}(\underline{x}-\underline{\mu})}$$

כאשר $\underline{\mu} = (\mu_1, \mu_2, ..., \mu_m)^t$ הוא וקטור התוחלות של $\underline{X}$, ו- Σ היא מטריצת השונויות המשותפות של $\underline{X}$, שרכיביה הם $\sigma_{ij} = \mathrm{Cov}(X_i, X_k) = \sum_{j=1}^{n} a_{ij}\, a_{kj}$ לכל $i,k = 1,2,...,m$.

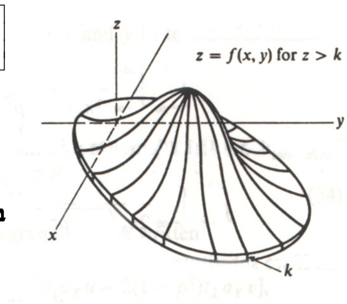
2. לכל X_i, המוגדר לעיל, יש התפלגות שולית נורמלית עם הפרמטרים μ_i ו- $\sum_{j=1}^{n} a_{ij}^2$.

3. לכל קבוצה חלקית של ה-X_i-ים, המוגדרים לעיל, יש התפלגות משותפת רב-נורמלית.

4. $M_{\underline{X}}(t_1,\ldots,t_m) = \exp\left\{\sum_{i=1}^{m} t_i\,\mu_i + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} t_i\, t_j \mathrm{Cov}(X_i, X_j)\right\}$ לכל t_1 , ... , t_m ממשיים.

אם ל-X ו-Y יש התפלגות משותפת רב-נורמלית, שנקראת במקרה הדו-ממדי **התפלגות דו-נורמלית**, אז הצפיפות המשותפת שלהם היא:

$$f_{X,Y}(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{(1-\rho^2)}}\cdot e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$



באיור: פונקציית צפיפות דו-נורמלית (בתחום שבו $f_{X,Y}(x,y)>k$)

טענות: אם ל-X ו-Y יש התפלגות משותפת דו-נורמלית, אז –

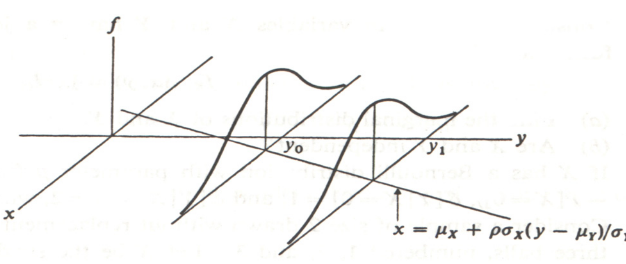
1. הפרמטר ρ, המופיע בפונקציית הצפיפות המשותפת שלהם, הוא מקדם המתאם ביניהם.

2. X ו-Y בלתי-תלויים אם ורק אם $\rho=0$, כלומר אם ורק אם הם בלתי-מתואמים.

3. ל-X בהינתן $Y=y$ יש התפלגות נורמלית עם תוחלת $\mu_X+\rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)=\mu_X+\frac{\text{Cov}(X,Y)}{\sigma_Y^2}(y-\mu_Y)$

ושונות $\sigma_X^2(1-\rho^2)=\left(\sigma_X^2-\frac{\text{Cov}(X,Y)^2}{\sigma_Y^2}\right)$.

באיור:
פונקציית צפיפות נורמלית מותנית של X בהינתן $Y=y$, עבור הערכים y_0 ו- y_1.
שימו לב, ש<u>התוחלת</u> של ההתפלגות המותנית היא פונקציה לינארית של y ואילו <u>שונותה</u> קבועה ביחס ל- y.
לכן, לכל ה"פעמונים" קו-מתאר דומה (הנקבע על-ידי השונות) ומיקום שונה (הנקבע על-ידי התוחלת).



יהי $X_1, X_2, \ldots, X_n$ **מדגם מקרי** מהתפלגות בעלת תוחלת μ ושונות σ^2.

ממוצע המדגם הוא $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$, ומתקיים: $E[\bar{X}]=\mu$ ו- $\text{Var}(\bar{X})=\sigma^2/n$

שונות המדגם היא $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$, ומתקיים: $E[S^2]=\sigma^2$

טענה: אם $X_1, X_2, \ldots, X_n$ הוא מדגם מקרי מהתפלגות נורמלית עם תוחלת μ ושונות σ^2, אז –

1. ממוצע המדגם $\bar{X}$ ושונות המדגם S^2 בלתי-תלויים זה בזה;

2. לממוצע המדגם $\bar{X}$ יש התפלגות נורמלית עם תוחלת μ ושונות σ^2/n;

3. למשתנה המקרי $(n-1)S^2/\sigma^2$ יש התפלגות חי-בריבוע עם $n-1$ דרגות-חופש.