Matematikai statisztika
I. témakör:
Valószínűségszámítási ismétlés

Elek Péter

1. Valószínűségi változók és eloszlások

1.1. Egyváltozós eset

Ismétlés: valószínűség fogalma

- Valószínűségekre vonatkozó axiómák
- Leszámlálási technikák
- Feltételes valószínűség
- Bayes-tétel
- Események függetlensége

Valószínűségi változók (v.v.)

- "Definíció":
  - formálisan: valószínűségi mezőn értelmezett mérhető függvény
  - nem formálisan: olyan változó, amely felvett értékét egy véletlen "kísérlet" kimenetele határozza meg
- Jelölés: X a v.v., x az általa felvett érték.
- Példa: érme feldobása 10-szer, és a fejek számának meghatározása
- Típusok: diszkrét és folytonos valószínűségi változók

Diszkrét valószínűségi változók

- véges vagy megszámlálhatóan végtelen számú lehetséges értéket vehet fel: $x_1, x_2, \ldots, x_k, \ldots$
- Példák: fej vagy írás érme feldobásakor, kockadobás kimenetele
- Valószínűség-eloszlás a lehetséges kimenetelek valószínűségeit adja meg: $\Pr(X = x_j) = p_j$, ahol $\sum p_j = 1$
  - Kockadobás: $\Pr(X = i) = 1/6$ $(i = 1, \ldots, 6)$
  - Érmedobás: $\Pr(X = \text{fej}) = \Pr(X = \text{írás}) = 1/2$

Diszkrét eloszlások: példák

- Bernoulli (p)
  - $\Pr(X = 1) = p$ és $\Pr(X = 0) = 1 - p$
- Binomiális (n, p)
  - n darab független, azonos Bernoulli-eloszlású v.v. összege: $Y = X_1 + X_2 + \cdot + X_n$, ahol X_i független Bernoulli (p) v.v.
  - $\Pr(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$ (ha $k = 0, 1, \ldots, n$)
  - Példa: fejek száma tíz érmedobás alatt
- Poisson (λ)
  - $\Pr(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ (ha $k \geq 0$ egész)
  - Példa: készülék meghibásodásainak száma (balesetek száma) egy adott évben ilyen lehet
- Geometriai (p)
  - $\Pr(X = k) = p(1-p)^{k-1}$ (ha $k \geq 1$ egész)
  - Példa: fej első bekövetkezéséhez szükséges érmedobások száma

Folytonos valószínűségi változók

- Nem diszkrét v.v.: megszámlálhatatlanul végtelen sok lehetséges értéket vehet fel (pl. $x \in \Re$, $x \in \Re^+$, $x \in [0, 1]$)
- (Abszolút) folytonos v.v. (nem formális definíció): minden konkrét értéket zérus valószínűséggel vesz fel
- Példa: egyenletes eloszlás a $[0, 1]$ intervallumon

- Minden részintervallumra annak hosszával arányos valószínűséggel esik.

- Példa részben nem folytonos és nem diszkrét valószínűségi változóra:
  - a zérus érték (p valószínűséggel) és egy egyenletes eloszlású változó ($1 - p$ valószínűséggel) keveréke

Eloszlásfüggvény

- Eloszlásfüggvény (cumulative distribution function, cdf): $F(x) = \Pr(X \leq x)$
  - annak valószínűsége, hogy X az x-nél nem nagyobb értéket vesz fel
- Minden (diszkrét vagy nem diszkrét) valószínűségi változóra értelmezhető.
- Tulajdonságok:
  - $\lim_{x\to-\infty} F(x) = 0$ és $\lim_{x\to\infty} F(x) = 1$
  - $F(x)$ nem csökkenő
  - $F(x)$ jobbról folytonos
  - $\Pr(X > a) = 1 - F(a)$
  - $\Pr(a < X \leq b) = \Pr(X \leq b) - \Pr(X \leq a) = F(b) - F(a)$

Folytonos valószínűségi változó sűrűségfüggvénye

- Sűrűségfüggvény (probability density function, pdf): $f(x) = F'(x)$, tehát az eloszlásfüggvény deriváltja
- Integrálásával megkapható annak valószínűsége, hogy X egy adott intervallumba esik: $\Pr(a < X \leq b) = \int_a^b f(x)dx$.
- Megjegyzés: $f(x)$ egyetlen konkrét érték valószínűségét sem mutatja, mert a folytonos esetben $\Pr(X = x) = 0$ minden x-re.
- Azonban $f(x)$ mégis megmutatja a "tipikus" értékeket: $\Pr(x < X \leq x + dx) \approx f(x)dx$ ha dx kicsi.
- Tulajdonságok:
  - $f(x) \geq 0$
  - $\int_{-\infty}^{\infty} f(x)dx = 1$

Folytonos eloszlások: példák

- Egyenletes: $U(a,b)$
  - $[a,b]$ intervallumon vesz fel értéket, minden részintervallumon annak hosszával arányos valószínűséggel
  - $f(x) = 1/(b-a)$ ha $a \leq x < b$ és zérus egyébként
- Normális: $N(\mu, \sigma^2)$
  - $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, ahol μ és $\sigma > 0$ paraméterek
- Exponenciális: $Exp(\theta)$
  - $f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$ ha $x > 0$ és zérus egyébként ($\theta > 0$ paraméter)
- Lognormális és Gamma-eloszlás (lásd később)

Valószínűségi változók transzformációja

- Legyen X egy v.v. és $g()$ egy valós értékű függvény. Ekkor $Y = g(X)$ is egy v.v.
- Legyen az X v.v. eloszlásfüggvénye $F_X(x)$ és sűrűségfüggvénye $f_X(x) = F'(x)$. Legyen $Y = g(X)$, ahol $g()$ szigorúan monoton növekvő. Ekkor Y eloszlás- és sűrűségfüggvénye a következő:
  - $F_Y(y) = F_X\left(g^{-1}(y)\right)$
  - $f_Y(y) = \left(g^{-1}\right)'(y) \times f_X\left[g^{-1}(y)\right]$.
- Bizonyítás: minden a esetén $F_X(a) = \Pr(X \leq a)$. Ekkor $F_Y(a) = \Pr(Y \leq a) = \Pr(g(X) \leq a) = \Pr\left(X \leq g^{-1}(a)\right) = F_X\left[g^{-1}(a)\right]$ (ahol felhasználtuk $g()$ szigorú monotonitását). Ezután $f_Y(a)$ deriválással megkapható.

Eloszlások szimulációja

- Hasznos állítás: Legyen U egyenletes eloszlású a $[0,1]$-en, $F()$ pedig folytonos eloszlásfüggvény. Ekkor az $F^{-1}(U)$ valószínűségi változó eloszlásfüggvénye éppen $F()$, ahol F^{-1} az F inverzfüggvénye.
- Bizonyítás: a feladatok között

1.2. Többváltozós eset

Együttes eloszlás (joint distribution)

- X és Y diszkrét v.v.-k együttes eloszlása: $\Pr(X = x_i, Y = y_j)$, ahol x_i és y_j a változók értékkészletének elemei.
  - Szokásos jelölés: $\Pr(X = x_i, Y = y_j) = p_{ij}$, ahol $\sum_i \sum_j p_{ij} = 1$.
- X és Y együttes eloszlásfüggvénye: $F_{X,Y}(x, y) = \Pr(X \leq x, Y \leq y)$
- X és Y folytonos v.v.-k együttes sűrűségfüggvénye: $f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$
  - $\Pr(x < X < x + dx, y < Y < y + dy) \approx f_{X,Y}(x, y)\, dxdy$, ha dx és dy kicsi.
- Hasonlóan kettőnél több v.v. esetén

Függetlenség

- X és Y v.v.-k *függetlenek*, ha az egyik értékének ismerete nem változtatja meg a másik eloszlását.
- Ekvivalens definíciók:
  - Együttes eloszlásfüggvénnyel: $F_{X,Y}(x, y) = F_X(x) F_Y(y)$ minden x, y esetén
  - Diszkrét v.v.-k esetén együttes eloszlással: $\Pr(X = x_i, Y = y_j) = \Pr(X = x_i) \Pr(Y = y_j)$ minden x_i, y_j esetén
  - Folytonos v.v.-k esetén együttes sűrűségfüggvénnyel: $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ minden x, y esetén
- Hasonlóan kettőnél több v.v. esetén

Példa

- Legyen X és Y együttes eloszlása a következő:

	y_1	y_2	$p_{i\cdot}$
x_1	0.4	0.3	0.7
x_2	0.2	0.1	0.3
$p_{\cdot j}$	0.6	0.4	1.0

- Ekkor a peremeloszlások (marginal distributions):
  - $\Pr(X = x_1) = 0.7$, $\Pr(X = x_2) = 0.3$

- $\Pr(Y = y_1) = 0.6$, $\Pr(Y = y_2) = 0.4$

- Független-e X és Y?
- Módosítsuk az együttes eloszlást úgy, hogy a két változó független legyen (változatlan peremeloszlásokkal)!

Feltételes eloszlás (conditional distribution)

- Y v.v. feltételes eloszlása X v.v.-ra vonatkozóan: Y eloszlása, ha tudjuk X értékét
- Diszkrét v.v. feltételes eloszlása: $\Pr(Y = y_j \mid X = x_i)$
  - Megjegyzés: $\Pr(Y = y_j \mid X = x_i) = \Pr\left(X = x_i, Y = y_j\right) / \Pr\left(X = x_i\right)$
- Folytonos v.v. feltételes sűrűségfüggvénye: $f_{Y|X}\left(y|x\right) = f_{X,Y}\left(x, y\right) / f_X\left(x\right)$
- X és Y akkor és csak akkor független, ha Y-nak X-re vonatkozó feltételes eloszlása (sűrűségfüggvénye) nem függ X-től: $f_{Y|X}(y \mid x) = f_{X,Y}(x, y)/f_X(x) = f_X(x) \cdot f_Y(y)/f_X(x) = f_Y(y)$.
- Példa (folyt.): az előző példában
  - Y feltételes eloszlása $X = x_1$ esetén: $\Pr(Y = y_1|X = x_1) = 4/7$, $\Pr(Y = y_2|X = x_1) = 3/7$.
  - Y feltételes eloszlása $X = x_2$ esetén: $\Pr(Y = y_1|X = x_2) = 2/3$, $\Pr(Y = y_2|X = x_2) = 1/3$.

2. Valószínűségi változók jellemzői

2.1. Középértékek

Középértékek

- Az eloszlás "tipikus" elemét mutatják meg.
- Várható érték
- Medián (folytonos v.v. esetén): az az m érték, amelyre $F(m) = 1/2$
  - A v.v. 50% valószínűséggel ennél kisebb, 50% valószínűséggel ennél nagyobb értéket vesz fel.
  - Diszkrét v.v. esetén nem feltétlenül egyértelmű.
- Módusz: diszkrét v.v. eloszlásának illetve folytonos v.v. sűrűségfüggvényének maximumhelye

- Nem feltétlenül egyértelmű.

- Szimmetrikus eloszlás esetén mindhárom mutató ugyanazt adja (ha létezik és egyértelmű).

Várható érték

- Diszkrét v.v.: $E(X) = \sum_{i=1}^{\infty} p_i x_i$
- Folytonos v.v.: $E(X) = \int_{-\infty}^{\infty} x f(x) dx$
- Tulajdonságok:
  - Konstans várható értéke: $E(c) = c$
  - a, b konstansok és X v.v. esetén $E(aX + b) = aE(X) + b$
  - Összeg (illetve lineáris kombináció) várható értéke: minden a_i valós szám és X_i v.v. $(i = 1, \ldots, n)$ esetén $E(a_1 X_1 + \ldots + a_n X_n) = a_1 E(X_1) + \ldots + a_n E(X_n)$.
- A várható érték minimalizálja a várható négyzetes eltérést: $h(b) = E\left((X - b)^2\right)$ minimumhelye $b = E(X)$.

Példák

- Bernoulli (p): $E(X) = p * 1 + (1 - p) * 0 = p$
- Binomiális (n, p): $E(X) = np$ (miért?)
- Normális (μ, σ^2): $E(X) = \mu$ (miért?)
- Exponenciális (θ): $E(X) = \theta$ (miért)
- Mi ezen eloszlások mediánja illetve módusza?

Valószínűségi változók függvényének várható értéke

- Legyen X v.v. és $g()$ egy valós értékű függvény. Ekkor $E(g(X))$ számításához nem feltétlenül szükséges $g(X)$ eloszlását meghatározni, mert
- $E(g(X)) = \sum_{i=1}^{\infty} p_i g(x_i)$ (diszkrét esetben) és
- $E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$ (folytonos esetben).
- Megjegyzés: nemlineáris g függvény esetén $E(g(X)) \neq g(E(X))$.
- Jensen-egyenlőtlenség: ha $g()$ konvex, akkor $E(g(X)) \geq g(E(X))$.
- Példa: határozzuk meg $E(X^2)$ értékét a kockadobásos példában!

2.2. Szóródási mutatók

Szóródási mutatók

- Terjedelem: a legnagyobb és legkisebb (lehetséges) érték különbsége
- Várható abszolút eltérés (a várható értéktől): $E\left(|X - E(X)|\right)$
- Variancia (szórásnégyzet): várható négyzetes eltérés a várható értéktől: $Var(X) = E\left[(X - E(X))^2\right] = E\left(X^2\right) - (E(X))^2$
  - Diszkrét v.v. esetén $E(X^2) = \sum_{i=1}^{k} p_i x_i^2$
  - Folytonos v.v. esetén $E\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f(x) dx$.
- Szórás: variancia négyzetgyöke: $\sigma_X = \sqrt{E\left(X^2\right) - (E(X))^2}$

Variancia és szórás tulajdonságai

- a, b konstansok és X v.v. esetén $Var(aX + b) = a^2 Var(X)$ és $\sigma_{aX+b} = |a|\,\sigma_X$.
- Konstans varianciája: $Var\,(c) = 0$.
- Valószínűségi változó standardizálása: legyen $E\,(X) = \mu$ and $Var\,(X) = \sigma^2$. Ekkor $Z = \frac{X-\mu}{\sigma}$ a standardizált v.v., amelyre $E\,(Z) = 0$ és $\sigma\,(Z) = 1$.

Példák

- Kockadobás varianciája: $\frac{35}{12}$
- Bernoulli (p) változó varianciája: $p(1 - p)$
- Binomial (n, p) változó varianciája: $np(1 - p)$
- Standard normális (azaz $\mu = 0$, $\sigma = 1$) változó varianciája: $Var(X) = 1$
- $N(\mu, \sigma^2)$ változó varianciája: $Var(X) = \sigma^2$.

2.3. Egyéb jellemzők

Eloszlások magasabb momentumai

- n-dik momentum: $E\left(X^n\right)$
- n-dik centrális momentum: $E\left\{[X - E\,(X)]^n\right\}$
- Variancia = második centrális momentum

- Ferdeség (standardizált harmadik centrális momentum) = $E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$
  - Eloszlás aszimmetriáját méri.
- Csúcsosság (kurtózis) (standardizált negyedik momentum) = $E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$
  - A normális eloszlás csúcsossága 3.
  - Az eloszlás csúcsosságát és vastag szélűségét méri.

Kvantilisek

- p-kvantilis (folytonos v.v. esetén): az a q érték, amelyre $F_X(q) = p$
- Kvartilisek (negyedek)
  - Az eloszlás az alsó kvartilisnél kisebb értéket 25%, a felső kvartilisnél kisebb értéket pedig 75% valószínűséggel vesz fel.
  - Második kvartilis = medián
- Decilisek (tizedek)
  - Példa: az eloszlás a harmadik decilisnél kisebb értéket 30%, annál nagyobb értéket pedig 70% valószínűséggel vesz fel
- Percentilisek (századok)

2.4. Többváltozós eloszlások jellemzői

Kovariancia és korreláció

- Kovariancia: $Cov(X,Y) = \sigma_{XY} = E\left[(X - E(X))(Y - E(Y))\right]$
  - Ekvivalens felírás: $Cov(X,Y) = E(XY) - E(X)E(Y)$
- Korreláció: $Corr(X,Y) = \frac{Cov(X,Y)}{sd(X) \cdot sd(Y)} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

Kovariancia és korreláció tulajdonságai

- Ha X és Y függetlenek, akkor $Cov(X,Y) = 0$
  - Bizonyítás: a várható érték tulajdonságai alapján
  - A megfordítás nem igaz: $\Pr(X = 1) = 0.25$, $\Pr(X = 0) = 0.5$, $\Pr(X = -1) = 0.25$, és legyen $Y = X^2$.

- Ha a, b, c, d konstansok és X, Y v.v.-k, akkor $Cov(aX + b, cY + d) = ac \cdot Cov(X, Y)$.
- Cauchy-Schwartz-egyenlőtlenség: $|Cov(X, Y)| \leq \sigma_X \cdot \sigma_Y$
- $-1 \leq Corr(X, Y) \leq 1$ (Cauchy-Schwartz-ból következik)
  - $Corr(X, Y) = 1$ tökéletes pozitív lineáris kapcsolat esetén
  - $Corr(X, Y) = -1$ tökéletes negatív lineáris kapcsolat esetén
- Korreláció az átskálázásra invariáns: $Corr(aX + b, cY + d) = sign(ac) \cdot Corr(X, Y)$.

Variancia további tulajdonságai

- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab \cdot Cov(X, Y)$
- Tehát $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$ akkor és csak akkor, ha X and Y korrelálatlanok.
- Az $\{X_1, \ldots, X_n\}$ páronként korrelálatlan v.v.-k és $\{a_1, \ldots, a_n\}$ valós számok esetén $Var\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 Var\left(X_i\right)$.

Feltételes várható érték

- $E(Y \mid X = x)$ az Y várható értéke abban az esetben, ha X egy konkrét x értéket vesz fel. Ez x függvénye, és megmutatja, hogy Y várható értéke hogyan függ X értékeitől.
- $E\left(Y \mid X = x\right) = \sum_{j=1}^{\infty} y_j \Pr\left(Y = y_j \mid X = x\right)$ (diszkrét esetben)
- $E\left(Y \mid X = x\right) = \int y f_{Y|X}\left(y \mid x\right) dy$ (folytonos esetben)
- Példa: X az iskolai évek száma és Y a havi bér. Ekkor $E(Y \mid X = 12)$ minden bizonnyal magasabb, mint $E(Y \mid X = 6)$.
- Megjegyzés: az $E\left(Y \mid X\right)$ jelölést arra a *valószínűségi változóra* használjuk, amely $E\left(Y \mid X = x\right)$ értéket vesz fel $X = x$ esetén. Ez tehát az X v.v. függvénye.

Feltételes várható érték tulajdonságai

- Minden $c()$ függvény esetén $E\left[c(X) \mid X\right] = c(X)$.
- Minden $a()$ és $b()$ függvény esetén $E\left[a(X)Y + b(X) \mid X\right] = a(X)E(Y \mid X) + b(X)$

- Ha X és Y függetlenek, akkor $E(Y \mid X) = E(Y)$.
  - Következmény: ha $E(U) = 0$ és U független X-től, akkor $E(U \mid X) = 0$.
- Minimalizáló tulajdonság:
  - Legyen $\mu(X) = E(Y \mid X)$. Ekkor az X változó minden $g(X)$ függvénye esetén
  - $E\left[(Y - \mu(X))^2 \mid X\right] \leq E\left[(Y - g(X))^2 \mid X\right]$ és
  - $E\left[(Y - \mu(X))^2\right] \leq E\left[(Y - g(X))^2\right]$.

Teljes várható érték tétele

- Teljes várható érték tétele: $E[E(Y \mid X)] = E(Y)$.
- Általánosítás: $E[E(Y \mid X)] = E[E(Y \mid X, Z) \mid X]$.
- Következmény: ha $E(Y \mid X) = E(Y)$, akkor $Cov(X, Y) = 0$. Sőt ilyenkor$X$ minden függvénye korrelálatlan Y-nal.
  - A megfordítás nem igaz. Ellenpélda?

Feltételes variancia

- Hasonlóan a feltétel nélküli esethez
- $Var(Y \mid X = x) = E\left(Y^2 \mid X = x\right) - (E(Y \mid X = x))^2$
- Hasznos tulajdonság:
  - Ha X és Y függetlenek, akkor $Var(Y \mid X = x) = Var(Y)$ minden x esetén.

Összefoglalás: függetlenség és hasonló fogalmak

- Függetlenség (independence): $f(Y \mid X = x)$ nem függ x-től
- "Mean independence": $E(Y \mid X = x)$ nem függ x-től
- Korrelálatlanság: $Corr(X, Y) = \sigma_{XY} = 0$
- Ezek fentről lefelé következnek egymásból, de fordítva nem.

Összefoglalás: néhány gyakran használt eloszlás jellemzői

Diszkrét	$\Pr(X=k)$		$E(X)$	$\sigma(X)$
Bernoulli	$p; 1-p$		p	$\sqrt{p(1-p)}$
Binomiális	$\binom{n}{k}p^k(1-p)^{n-k}$		np	$\sqrt{np(1-p)}$
Geometriai	$p(1-p)^{k-1}$		$1/p$	$\sqrt{1-p}/p$
Poisson	$\frac{\lambda^k}{k!}e^{-\lambda}$		λ	$\sqrt{\lambda}$
Folytonos	$f(x)$	$F(x)$	$E(X)$	$sd(X)$
Egyenletes	$\frac{1}{b-a}$	$\frac{x-a}{b-a}$	$\frac{a+b}{2}$	$\frac{b-a}{\sqrt{12}}$
Normális	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$	μ	σ
Exponenciális	$\frac{1}{\theta}e^{-\frac{x}{\theta}}$	$1-e^{-\frac{x}{\theta}}$	θ	θ

3. A normális eloszlás és kapcsolódó eloszlások

3.1. A normális eloszlás és tulajdonságai

A normális eloszlás fontos, mert

- Számos valóságos változó normális eloszlású (például férfiak illetve nők testmagassága)
- Aszimptotika: számos (szokásos) becslőfüggvény aszimptotikusan normális eloszlású (lásd később)
- Számos gyakran használt eloszlás származik belőle: lognormális, χ^2-, t- és F-eloszlás

Kétváltozós normális eloszlás

- Kétváltozós standard normális eloszlás sűrűségfüggvénye:

$$f(x,y) = \frac{1}{2\pi}\exp\left[-\frac{x^2+y^2}{2}\right]$$

  - Peremeloszlások független standard normális eloszlások

- Kétváltozós normális eloszlás sűrűségfüggvénye, ha a peremeloszlások ρ korrelációjú standard normális eloszlások:

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left[-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right]$$

- Ez tovább transzformálható a várható értékekkel és a szórásokkal.

A normális eloszlás tulajdonságai

- Ha $X \sim N(\mu,\sigma^2)$, akkor $aX+b \sim N(a\mu+b,(|a|\sigma)^2)$.

- Tehát ha $X \sim N(\mu, \sigma^2)$, akkor $(X - \mu)/\sigma \sim N(0, 1)$ (standard normális eloszlás).
- Ha X és Y együttes eloszlása normális, akkor a függetlenség és a korrelálatlanság ekvivalens.
- Legyen $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$ és a két változó legyen független egymástól. Ekkor $X + Y \sim N\left(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2\right)$.
- Legyen $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$, és a két változó legyen független egymástól. Ekkor a és b konstansok esetén
$$aX + bY \sim N\left(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2\right).$$

Normális eloszlásra vonatkozó valószínűségek számítása

- A $\Phi(z)$ standard normális eloszlásfüggvény nem határozható meg zárt alakban, de természetesen numerikusan közelíthető.
- A normális eloszlásra vonatkozó valószínűségek a standard normális eloszlásfüggvény segítségével kifejezhetőek.
- Legyen $X \sim N\left(\mu, \sigma^2\right)$. Ekkor $\Pr(a < X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$.
- Példa: $\Pr\left(|\frac{X-\mu}{\sigma}| \leq 1.96\right) = \Phi(1.96) - \Phi(-1.96) = 2 \cdot \Phi(1.96) - 1 = 0.95$

3.2. Kapcsolódó eloszlások

Lognormális eloszlás

- Ha $Y \sim N(\mu, \sigma^2)$, akkor $X = e^Y$ lognormális eloszlású.
- Tulajdonságok:
  - $E\left(e^Y\right) = e^{\mu + \frac{\sigma^2}{2}}$
  - $Var\left(e^Y\right) = \sigma^2 e^{2\mu + \sigma^2}$

χ^2 eloszlás

- Ha $Z_1, Z_2, \ldots, Z_n$ *független standard normális eloszlású* v.v.-k, akkor $X = \sum_{i=1}^n Z_i^2$ egy χ_n^2 eloszlású v.v. (ahol n a szabadsági fok).
- Tulajdonságok:
  - $E(X) = n$
  - $Var(X) = 2n$ (miért?)
  - Csak pozitív értékeket vesz fel

t-eloszlás

- Ha $Z \sim N(0,1)$, $X \sim \chi^2_n$ és függetlenek egymástól, akkor $t = \frac{Z}{\sqrt{X/n}}$ t-eloszlást követ n szabadsági fokkal.
- Az eloszlás alakja hasonló a standard normális eloszláséhoz: szimmetrikus de vastagabb szélű (azaz szélsőséges megfigyelések nagyobb valószínűséggel következnek be).
- Várható érték zérus (ha $n > 1$), variancia $\frac{n}{n-2}$ (ha $n > 2$) (egyébként ezek a momentumok nem léteznek)
- Ha $n \to \infty$, akkor a t-eloszlás tart a normális eloszláshoz. (Bizonyítás: lásd később)

F-eloszlás

- Ha $X_1 \sim \chi^2_{k_1}$, $X_2 \sim \chi^2_{k_2}$és egymástól függetlenek, akkor $F = \frac{X_1/k_1}{X_2/k_2}$ egy F_{k_1,k_2}-eloszlást követ (ahol k_1 és k_2 a szabadsági fokok).
- Csak pozitív értékeket vehet fel.
- $t_n^2 \sim F_{1,n}$.

Tananyag

- W Appendix B
- A 1-5 (kivéve a bizonyítások, Theorem 3.6.3, section 3.7, definition 5.3.1, section 5.4)