Néhány nevezetes eloszlás

Transzformált valószínűségi változók

Medvegyev Péter

2011

Egy dimenzióban az

$$\int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(y))\,g'(y)\,dy$$

helyettesítési formula a Newton–Leibniz formula közvetlen folyománya. Mivel többdimenzióban a formula nem érvényes, ezért némiképpen át kell írni. Ha a g monoton nő, akkor mivel $g([a,b]) = [g(a), g(b)]$, ezért

$$\int_{g([a,b])} f(x)\,dx = \int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(y))\left|g'(y)\right| dy,$$

ha monoton csökken, akkor $g([a,b]) = [g(b), g(a)]$

$$\int_{g([a,b])} f(x)\,dx = -\int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(y))\left|g'(y)\right| dy.$$

Alapelv, integráltranszformációs tétel

$\mathbb{R}^n$ esetén a második megfogalmazás használható.

Theorem

Legyen $U \subseteq \mathbb{R}^n$ egy nyílt halmaz és $T : U \to \mathbb{R}^n$ egy invertálható és folytonosan differenciálható leképezés. Ekkor tetszőleges $B \in \mathcal{B}(\mathbb{R}^n)$, $B \subseteq U$ esetén

$$\int_{T(B)} f(x)\, d\lambda_n(x) = \int_B f(T(y)) \left|\det\left(T'(y)\right)\right| d\lambda_n(y),$$

ahol $T'(y)$ a T parciális deriváltjaiból álló Jacobi-mátrix és ahol λ_n szokásos a téglatestek térfogata által generált úgynevezett Lebesgue-mérték.

Alapelv, integráltranszformációs tétel

A tételben a leginkább meglepő elem, hogy miként kerül az integráltranszformációs tételbe a Jacobi-determináns. Ezt a következő némiképpen elnagyolt gondolatmenettel indokolhatjuk: Jelölje $\mathrm{vol}\left(B\right)$ valamely $B \subseteq \mathbb{R}^n$ halmaz "térfogatát". Megmutatjuk, hogy ekkor tetszőleges $\mathbf{A}$ lineáris leképezésre igaz a következő:

$$\lambda_n\left(\mathbf{A}B\right) \stackrel{\circ}{=} \mathrm{vol}\left(\mathbf{A}B\right) = \left|\det\left(\mathbf{A}\right)\right| \mathrm{vol}\left(B\right) \stackrel{\circ}{=} \left|\det\left(\mathbf{A}\right)\right| \lambda_n\left(B\right).$$

Alapelv, integráltranszformációs tétel

A térfogat pontos tartalmát a mértékelmélet keretében szokás tisztázni, mi a következő feltételezésekkel élünk:

1. Az egységkocka térfogata 1, illetve tetszőleges téglatest térfogata az oldalak hosszának szorzata.
2. A térfogat invariáns a forgatásra és az eltolásra nézve.
3. A térfogat homogén abban az értelemben, hogy tetszőleges $\alpha > 0$ szám esetén ha valamelyik koordinátát α-val megszorzunk, akkor a térfogat is α-val nő. Így többek között minden $\alpha > 0$ esetén $\mathrm{vol}\left(\alpha B\right) = \alpha^n \mathrm{vol}\left(B\right)$, illetve, ha Λ egy diagonális mátrix, akkor $\mathrm{vol}\left(\Lambda B\right) = \lambda_1 \lambda_2 \cdots \lambda_n \mathrm{vol}\left(B\right)$, ahol λ_i a Λ diagonálisában levő elemek.

Alapelv, integráltranszformációs tétel

Egyedül az utolsó tulajdonság nem tűnik evidensnek. Indoklása a következő: Nyilván a tulajdonság igaz koordináta irányú téglatestekre. Ebből következően igaz véges számú ilyen téglatest diszjunkt egyesítésére is. Ezeket a halmazokat szokás elemi halmazoknak nevezni. Elemi halmazok, metszete, egyesítése, illetve komplementere elemi halmaz, vagyis az elemi halmazok algebrát alkotnak, de nem alkotnak σ-algebrát. Ahhoz, hogy a térfogat használható fogalom legyen ki kell terjeszteni. Szemléletesen a térfogat fogalma olyan halmazokra használható, amelyek valamiképpen elemi halmazokkal jól közelíthetőek. Ilyen közelítés például ha a halmaz elemi halmazok megszámlálható sorozatának monoton csökkenő metszete, vagy monoton növekedő egyesítése. A pontos részletek elmagyarázása messze vezetne, illetve a mértékelméleti tárgyalást igényel. (Egész pontosan a mértékelmélet egyik célja pontosan annak indoklása, hogy a térfogat fogalma egyértelműen kiterjeszthető egy az elemi halmazokat tartalmazó σ-algebrára.)

Alapelv, integráltranszformációs tétel

A közvetlen szemlélet alapján is viszonylag könnyen elfogadható azonban, hogy a "szokásos" halmazok jól közelíthetőek elemi halmazokkal, így a térfogat fogalma, a fenti tulajdonságokkal alkalmazható rájuk.
Érdemes megjegyezni, hogy a forgatásinvariancia gömbök esetén nyilvánvaló, a homogenitás pedig koordináta irányú téglatestek esetén nyilvánvaló. A fenti tulajdonságok indoklásához többek között azt is indokolni kell, hogy a téglatestek elforgatásakor a térfogatuk nem változik, vagy hogy a gömbök koordináta irányú nyújtásakor a térfogat a fenti homogén módon alakul. Vagyis, hogy a térfogat két alaptulajdonsága más és más "elemi" halmazok esetén nyilvánvaló, és amit indokolni kell, az az, hogy egyrészt van közös általánosítás, másrészt a megfelelő tulajdonságok egyszerre öröklődnek.

Alapelv, integráltranszformációs tétel

A determinánsos képlet triviálisan teljesül, ha az $\mathbf{A}$ oszlopai összefüggnek, ugyanis ilyenkor a determináns nulla, és mivel az $\mathbf{A}B$ az $\mathbb{R}^n$ egy valódi alterébe esik

$$\mathrm{vol}\left(\mathbf{A}B\right) = 0 = \left|\det\left(\mathbf{A}\right)\right| \mathrm{vol}\left(B\right).$$

Az általános eset a komplex számok $z = r\exp\left(i\varphi\right)$ felbontásának általánosításának tekinthető poláris felbontás következménye:

Lemma

Ha $\mathbf{A}$ egy invertálható mátrix, akkor az $\mathbf{A}$ felbontható $\mathbf{A} = \mathbf{RU}$ alakba, ahol az $\mathbf{R}$ pozitív definit, az $\mathbf{U}$ pedig ortonormált.

Alapelv, integráltranszformációs tétel

Az $\mathbf{AA}^T$ mátrix szimmetrikus és pozitív definit, így $\mathbf{O}\Lambda\mathbf{O}^T$ alakba írható, ahol az $\mathbf{O}$ az $\mathbf{AA}^T$ sajátvektoraiból álló ortonormált mátrix, Λ a sajátértékeket tartalmazó diagonális mátrix. Mivel a $\Lambda \geq 0$, ezért az $\mathbf{R} \circeq \mathbf{O}\sqrt{\Lambda}\mathbf{O}^T$ definíció értelmes. Értelemszerűen a gyökjel a Λ elemeire értendő. Az $\mathbf{R}$ invertálható és elegendő igazolni, hogy az $\mathbf{U} \circeq \mathbf{R}^{-1}\mathbf{A}$ mátrix ortonormált. Szimmetrikus mátrix inverze is szimmetrikus, így

$$
\begin{aligned}
\mathbf{UU}^T &= \mathbf{R}^{-1}\mathbf{A}\left(\mathbf{R}^{-1}\mathbf{A}\right)^T = \mathbf{R}^{-1}\mathbf{AA}^T\left(\mathbf{R}^{-1}\right)^T = \mathbf{R}^{-1}\mathbf{AA}^T\mathbf{R}^{-1} = \\
&= \mathbf{R}^{-1}\mathbf{O}\Lambda\mathbf{O}^T\mathbf{R}^{-1} = \mathbf{R}^{-1}\mathbf{O}\sqrt{\Lambda}\mathbf{I}\sqrt{\Lambda}\mathbf{O}^T\mathbf{R}^{-1} = \\
&= \mathbf{R}^{-1}\mathbf{O}\sqrt{\Lambda}\left(\mathbf{O}^T\mathbf{O}\right)\sqrt{\Lambda}\mathbf{O}^T\mathbf{R}^{-1} \\
&= \left(\mathbf{R}^{-1}\mathbf{O}\sqrt{\Lambda}\mathbf{O}^T\right)\left(\mathbf{O}\sqrt{\Lambda}\mathbf{O}^T\mathbf{R}^{-1}\right) = \\
&= \left(\mathbf{R}^{-1}\mathbf{R}\right)\left(\mathbf{R}\mathbf{R}^{-1}\right) = \mathbf{I}.
\end{aligned}
$$

Alapelv, integráltranszformációs tétel

A poláris felbontással és a térfogat alaptulajdonságaival a determinánsos formula már egyszerűen igazolható. (Vegyük észre, hogy mivel az ortonormált mátrixok nem módosítják a skaláris szorzatot, ezért nem módosítják a vektorok hosszát sem, vagyis definíció szerint forgatások, így nem módosítják a térfogatot.) Ha az **A** invertálható, akkor

$$
\begin{aligned}
\operatorname{vol}(\mathbf{A}B) &= \operatorname{vol}(\mathbf{R}\mathbf{U}B) = \operatorname{vol}(\mathbf{R}(\mathbf{U}B)) = \\
&= \operatorname{vol}(\mathbf{R}B) = \operatorname{vol}\left(\mathbf{O}\sqrt{\Lambda}\mathbf{O}^T B\right) = \\
&= \operatorname{vol}\left(\mathbf{O}\sqrt{\Lambda}B\right) = \sqrt{\prod_{i=1}^{n}\lambda_i}\operatorname{vol}(\mathbf{O}B) = \sqrt{\prod_{i=1}^{n}\lambda_i}\operatorname{vol}(B) = \\
&= \sqrt{\det\left(\mathbf{R}\mathbf{R}^T\right)}\operatorname{vol}(B) = \sqrt{\det(\mathbf{R})\det\left(\mathbf{R}^T\right)}\operatorname{vol}(B) = \\
&= \sqrt{\det(\mathbf{R})\det(\mathbf{R})}\operatorname{vol}(B) = |\det(\mathbf{R})|\operatorname{vol}(B) = \det(\mathbf{R})\operatorname{vol}(B).
\end{aligned}
$$

Az $\mathbf{U}$ mátrix ortonormált, ezért

$$1 = \det(\mathbf{I}) = \det\left(\mathbf{U}\mathbf{U}^T\right) = \det(\mathbf{U})\det\left(\mathbf{U}^T\right) = \det(\mathbf{U})^2$$

így $\det(\mathbf{U}) = \pm 1$. Ebből következően, felhasználva, hogy az $\mathbf{R}$ pozitív definit, így a determinánsa pozitív

$$\begin{aligned}
\mathrm{vol}(\mathbf{A}B) &= \left|\det(\mathbf{R})\det(\mathbf{U})\right|\mathrm{vol}(B) = \\
&= \left|\det(\mathbf{R}\mathbf{U})\right|\mathrm{vol}(B) = \left|\det(\mathbf{A})\right|\mathrm{vol}(B)
\end{aligned}$$

A determinánsos képlet alapján a tétel "indoklása" a következő: Tekintsük a B halmaz egy $\left(V_i^{(k)}\right)$ partícióját. Az invertálhatóság miatt ez a $T(B)$ halmazon egy $U_i^{(k)} \stackrel{\circ}{=} T\left(V_i^{(k)}\right)$ partíciót definiál..Az integrál definíciója szerint

$$\int_{T(B)} f(x)\,dx \approx \sum_i f(x_i)\operatorname{vol}\left(U_i^{(k)}\right) = \sum_i f(T(y_i))\operatorname{vol}\left(U_i^{(k)}\right),$$

ahol $x_i \in U_i^{(k)}$.és $x_i = T(y_i)$ és $y_i \in V_i^{(k)}$. A bizonyítás lényege, hogy ha a $V_i^{(k)}$ elég "kicsi", akkor a T a $V_i^{(k)}$ halmazon jól közelíthető a deriváltjával.

Alapelv, integráltranszformációs tétel

Ezért, felhasználva az eltolásinvarienciát

$$\begin{aligned}
\mathrm{vol}\left(U_i^{(k)}\right) = \mathrm{vol}\left(T\left(V_i^{(k)}\right)\right) = \\
= \mathrm{vol}\left(T\left(y_i\right) + T'\left(y_i\right)\left(V_i^{(k)} - y_i\right) + o\left(\left\| V_i^{(k)} - y_i \right\|\right)\right) \approx \\
\approx \mathrm{vol}\left(T\left(y_i\right) + T'\left(y_i\right)\left(V_i^{(k)} - y_i\right)\right) = \mathrm{vol}\left(T'\left(y_i\right)\left(V_i^{(k)}\right)\right) = \\
= \left|\det\left(T'\left(y_i\right)\right)\right| \mathrm{vol}\left(V_i^{(k)}\right).
\end{aligned}$$

Így

$$\begin{aligned}\int_{T(B)} f(x)\,dx &\approx \sum_i f(T(y_i))\left|\det\left(T'(y_i)\right)\right| \mathrm{vol}\left(V_i^{(k)}\right) \approx \\ &\approx \int_B f(T(y))\left|\det\left(T'(y)\right)\right| dy.\end{aligned}$$

Alapelv, integráltranszformációs tétel

A bizonyítás során garantálni kell egyrészt, hogy a $\left(V_i^{(k)}\right)$ és az $\left(U_i^{(k)}\right)$ partíciók egyszerre legyenek infinitezimálisak, ha $k \to \infty$, másrészt hogy a másodrendű $\mathrm{vol}\left(o\left(\left\|V_i^{(k)} - y_i\right\|\right)\right)$ hibák elhagyása során a hibák összege is nullához tart. Az első a folytonos deriválhatóságból következik, ugyanis a folytonosság miatt a deriváltak a kompakt halmazokon korlátosak, így alkalmas K-ra, legalább a kompakt halmazokon $\|T(x) - T(y)\| \leq K\|x - y\|$. A második tulajdonság abból következik, hogy ha a $V_i^{(k)}$ helyett $\alpha V_i^{(k)}$-t veszünk, például minden oldalt felezünk, akkor a térfogat $\alpha^n \mathrm{vol}\left(V_i^{(k)}\right)$ lesz. Ugyanakkor ilyenkor $1/\alpha^n$-szeresére nő a partícióban levő halmazok száma. Összességében tehát az összeg továbbra is nullához fog tartani.

Alapelv, integráltranszformációs tétel

Ha $(\xi_k)_{k=1}^n$ az $(\Omega, \mathcal{A}, \mathbf{P})$ téren értelmezett valószínűségi változók, és g az $\mathbb{R}^n$ téren értelmezett Borel-mérhető függvény, akkor az

$$\eta \circeq g(\xi_1, \ldots, \xi_n)$$

összefüggéssel definiált függvény szintén $\mathcal{A}$-mérhető, vagyis valószínűségi változó. Hogyan lehet meghatározni az η eloszlását?

Theorem

Ha a $(\xi_1, \xi_2, \ldots, \xi_n)$ vektor együttes sűrűségfüggvénye f, és valamely h függvényre minden $B \in \mathcal{B}(\mathbb{R})$ halmazra

$$\int_B h d\lambda = \int_{g^{-1}(B)} f d\lambda_n,$$

akkor a h éppen az $\eta \circeq g(\xi_1, \xi_2, \ldots, \xi_n)$ sűrűségfüggvénye.

Alapelv, integráltranszformációs tétel

Bizonyítás: Emlékeztetünk, hogy definíció szerint egy $u : \mathbb{R}^n \to \mathbb{R}_+$ valamely ζ vektorváltozó sűrűségfüggvénye, ha minden $B \in \mathcal{B}(\mathbb{R})$ esetén

$$\int_B u d\lambda_n = \mathbf{P}(\zeta \in B),$$

ahol λ_n jelöli az $\mathbb{R}^n$-en értelmezett szokásos mértéket, vagyis az úgynevezett Lebesgue-mértéket. Ha $B \in \mathcal{B}(\mathbb{R})$, akkor mivel f a $(\xi_1, \ldots, \xi_n)$ sűrűségfüggvénye, ezért

$$\begin{aligned}
\int_B h d\lambda &= \int_{g^{-1}(B)} f d\lambda_n = \\
&= \mathbf{P}\left((\xi_1, \xi_2, \ldots, \xi_n) \in g^{-1}(B)\right) = \\
&= \mathbf{P}\left(g(\xi_1, \xi_2, \ldots, \xi_n) \in B\right) \overset{\circ}{=} \mathbf{P}(\eta \in B).
\end{aligned}$$

Alapelv, integráltranszformációs tétel

Theorem

Ha f a $(\xi_1, \ldots, \xi_k)$ vektor sűrűségfüggvénye,

$$(\eta_1, \ldots, \eta_k) \mathrel{\mathring{=}} \varphi(\xi_1, \xi_2, \ldots, \xi_k),$$

és a φ^{-1} függvény létezik és differenciálható, akkor az $(\eta_1, \ldots, \eta_k)$ sűrűségfüggvénye

$$h(x_1, \ldots, x_k) \mathrel{\mathring{=}} f\left(\varphi^{-1}(x_1, \ldots, x_n)\right) \left| \det\left(\frac{d}{dx}\varphi^{-1}(x_1, \ldots, x_n)\right) \right|.$$

Alapelv, integráltranszformációs tétel

Bizonyítás: A többdimenziós integrál transzformációs tétel alapján minden B Borel-halmazra

$$\begin{aligned}
&\mathbf{P}\left(\left(\eta_1, \eta_2, \ldots \eta_n\right) \in B\right) \circeq \\
&\circeq \mathbf{P}\left(\left(\varphi_1\left(\xi_1, \ldots, \xi_n\right), \ldots, \varphi_n\left(\xi_1, \ldots, \xi_n\right)\right) \in B\right) \\
&= \mathbf{P}\left(\left(\xi_1, \ldots, \xi_n\right) \in \varphi^{-1}(B)\right) = \int_{\varphi^{-1}(B)} f d\lambda_n = \\
&= \int_{T(B)} f d\lambda_n = \int_B f(T)\left|\det\left(T'\right)\right| d\lambda_n,
\end{aligned}$$

feltéve persze, hogy a tétel feltételei teljesülnek, vagyis a $T \circeq \varphi^{-1}$ létezik és differenciálható. Vagyis $f(T)\left|\det\left(T'\right)\right|$ az $\left(\eta_k\right)$ vektor sűrűségfüggvénye.

Alapelv, integráltranszformációs tétel

Example

Határozzuk meg az összeg, a szorzat és a hányados sűrűségfüggvényét.

Összeg sűrűségfüggvénye

1. Összeg: Ilyenkor

$$\begin{aligned}\varphi(x,y) &= (x+y,y) \overset{\circ}{=} (u,v).\\ \varphi^{-1}(u,v) &= (u-v,v) = (x,y).\end{aligned}$$

Feltéve, hogy a ξ sűrűségfüggvénye h az η sűrűségfüggvénye g. A feltételezett függetlenség miatt a (ξ,η) sűrűségfüggvénye $f(x,y) = h(x)g(y)$. A $(\xi+\eta,\eta)$ sűrűségfüggvénye

$$\begin{aligned}f(T)\left|\det(T')\right| &= h(u-v)g(v)\left|\det\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}\right| =\\ &= h(u-v)g(v),\end{aligned}$$

A $\xi+\eta$ eloszlása az első peremeloszlás, vagyis a $\xi+\eta$ sűrűségfüggvénye

$$\int_{-\infty}^{\infty} h(u-v)g(v)\,dv.$$

Ha nem függetlenek, akkor az összeg sűrűségfüggvénye $\int_{-\infty}^{\infty} f(u-v,v)\,dv$

Összeg sűrűségfüggvénye

$w = u - v$ helyettesítéssel

$$\int_{-\infty}^{\infty} f\left(u - v, v\right) dv = \int_{\infty}^{-\infty} f\left(w, u - w\right)\left(-1\right) dw = \int_{-\infty}^{\infty} f\left(w, u - w\right) dw$$

amely egybevág avval, hogy az összeadás kommutatív.

Szorzat sűrűségfüggvénye

2. Szorzat. Ilyenkor

$$\begin{aligned}\varphi(x,y) &= (xy,y) \mathrel{\hat{=}} (u,v)\,.\\ \varphi^{-1}(u,v) &= (u/v,v) = (x,y)\,.\end{aligned}$$

$$\begin{aligned}f(T)\left|\det(T')\right| &= h(u/v)\,g(v)\left|\det\begin{pmatrix} 1/v & -u/v^2 \\ 0 & 1\end{pmatrix}\right| =\\ &= h(u/v)\,g(v)\,\frac{1}{|v|},\end{aligned}$$

A $\xi\eta$ eloszlása az első peremeloszlás, vagyis a $\xi\eta$ sűrűségfüggvénye

$$\int_{-\infty}^{\infty} h\left(\frac{u}{v}\right) g(v)\,\frac{1}{|v|}dv.$$

Ha nem függetlenek, akkor $\int_{-\infty}^{\infty} f\left(\frac{u}{v},v\right)\frac{1}{|v|}dv$

Szorzat sűrűségfüggvénye

$w = u/v$ helyettesítéssel

$$v = \frac{u}{w}, \frac{dv}{dw} = -\frac{u}{w^2}, \Rightarrow dv = -\frac{u}{w^2}dw$$

Ha az u pozitív, akkor a $(0,\infty)$ ráképződik a $(\infty,0)$-ra illetve a $(-\infty,0)$ a $(0,-\infty)$ re képződik. Ezért a határokat fel kell cserélni. Ilyenkor $u = |u|$ és ezért

$$\begin{aligned}\int_{-\infty}^{\infty} f\left(\frac{u}{v}, v\right)\frac{1}{|v|}dv &= -\int_{-\infty}^{\infty} f\left(w, \frac{u}{w}\right)\frac{|w|}{|u|}\left(-\frac{u}{w^2}\right)dw = \\ &= \int_{-\infty}^{\infty} f\left(w, \frac{u}{w}\right)\frac{1}{|w|}dw\end{aligned}$$

Ha az u negatív a $(0,\infty)$ ráképződik a $(-\infty,0)$-ra illetve a $(-\infty,0)$ a $(0,\infty)$ re képződik. Ilyenkor $|u| = -u$

$$\begin{aligned}\int_{-\infty}^{\infty} f\left(\frac{u}{v}, v\right)\frac{1}{|v|}dv &= \int_{-\infty}^{\infty} f\left(w, \frac{u}{w}\right)\frac{|w|}{|u|}\left(-\frac{u}{w^2}\right)dw = \\ &= \int^{\infty} f\left(w, \frac{u}{w}\right)\frac{1}{|w|}dw.\end{aligned}$$

Hányados sűrűségfüggvénye

3. Hányados. Ilyenkor

$$\begin{aligned}\varphi(x,y) &= (x/y,y) \overset{\circ}{=} (u,v).\\ \varphi^{-1}(u,v) &= (u\cdot v,v) = (x,y).\end{aligned}$$

$$\begin{aligned}f(T)\left|\det\left(T'\right)\right| &= h(u\cdot v)\,g(v)\left|\det\begin{pmatrix} v & u\\ 0 & 1\end{pmatrix}\right| =\\ &= h(u\cdot v)\,g(v)\,|v|,\end{aligned}$$

A ξ/η eloszlása az első peremeloszlás, vagyis a ξ/η sűrűségfüggvénye

$$\int_{-\infty}^{\infty} h(uv)\,g(v)\,|v|\,dv.$$

Ha nem függetlenek, akkor $\int_{-\infty}^{\infty} f(uv,v)\,|v|\,dv$.

Hányados sűrűségfüggvénye

$w = uv$ helyettesítéssel, ha u pozitív, akkor a határokat nem kell módosítani

$$\frac{dw}{dv} = u \Rightarrow dv = \frac{dw}{u}$$

$$\int_{-\infty}^{\infty} f\left(uv, v\right)\left|v\right| dv = \int_{-\infty}^{\infty} f\left(w, \frac{w}{u}\right)\left|\frac{w}{u}\right| \frac{dw}{u}$$

ami egybevág avval, hogy az osztás nem kommutatív.

Valószínűségi változók szimulálása

Ha a ξ eloszlása egyenletes a $[0,1]$ intervallumon és $\eta \mathrel{\hat{=}} F^{-1}(\xi)$, ahol F egy folytonos, szigorúan monoton növekedő eloszlásfüggvény, akkor

$$\mathbf{P}(\eta < x) = \mathbf{P}\left(F^{-1}(\xi) < x\right) = \mathbf{P}(\xi < F(x)) = F(x),$$

vagyis az η eloszlásfüggvénye F. Megfordítva, ha az η eloszlásfüggvénye F, és $0 < x < 1$, akkor

$$\mathbf{P}(F(\eta) < x) = \mathbf{P}\left(\eta < F^{-1}(x)\right) = F\left(F^{-1}(x)\right) = x,$$

vagyis ilyenkor az $F(\eta)$ eloszlása egyenletes.

Example

Exponenciális eloszlású változó generálása.

$F(x) = 1 - \exp(-\lambda x)$. Ebből

$$F^{-1}(y) = -\lambda^{-1} \ln(1-y).$$

Mivel ha ξ egyenletes a $[0,1]$-en, akkor az $1-\xi$ is az, így az

$$\eta \circeq -\lambda^{-1} \ln \xi$$

eloszlása λ paraméterű exponenciális.

Example

Cauchy-eloszlású változó szimulálása

Ha η egy Cauchy-eloszlású valószínűségi változó, akkor az η sűrűségfüggvénye

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

Ebből az eloszlásfüggvénye

$$\begin{aligned} F(x) &= \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+t^2} dt = [\arctan]_{-\infty}^{x} = \frac{1}{\pi}\left(\arctan x + \frac{\pi}{2}\right) = \\ &= \frac{1}{\pi}\arctan x + \frac{1}{2}, \\ F^{-1}(y) &= \tan \pi\left(y - \frac{1}{2}\right), \end{aligned}$$

vagyis ha a ξ egyenletes a $[0,1]$ szakaszon, akkor az $\eta \circeq \tan \pi(\xi - 1/2)$ Cauchy-eloszlású.

Example

Standard normális eloszlású valószínűségi változó szimulálása.

Az előzőek alapján kézenfekvőnek látszik, hogy az $N(0,1)$ változók szimulálására a $\Phi^{-1}(y)$ függvényt használjuk. Azonban ennek kiszámolása bonyolult, így ritkán használatos. Az itt bemutatott módszert szokás Box–Müller módszernek is mondani.

Theorem

Ha δ_1, δ_2 független, a $(0,1)$ intervallumon egyenletes eloszlású változók, akkor a

$$\xi \doteq \sqrt{-2\ln\delta_1}\cos(2\pi\delta_2), \quad \eta \doteq \sqrt{-2\ln\delta_1}\sin(2\pi\delta_2)$$

változók függetlenek, és $\xi \cong \eta \cong N(0,1)$.

Bizonyítás: A módszer igazolása céljából vegyünk két független $N(0,1)$ változót, és a (ξ, η) párt tekintsük az $\mathbb{R}^2$ sík véletlenül kiválasztott pontjának.

$$\begin{aligned} f(x,y) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) = \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left(x^2+y^2\right)\right). \end{aligned}$$

Mi történik, ha (ρ, φ) polárkoordinátákra térünk át? Legyen

$$T(\rho,\varphi) \stackrel{\circ}{=} \rho(\cos\varphi, \sin\varphi) = (\xi, \eta)$$

a polárkoordinátákról való visszatérést megadó inverz leképezés. A T az origón kívül, vagyis a $\rho > 0$ tartományon injektív, és az origótól eltekintve teljesíti az integráltranszformációs tétel feltételeit.

$$\left|\det\left(T'\right)\right| = \left|\det\begin{pmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{pmatrix}\right| = \rho.$$

Valószínűségi változók szimulálása

A már bemutatott módon a polárkoordinátákban a sűrűségfüggvény

$$s(r,t) = f(T)\left|\det(T')\right| = \frac{1}{2\pi}\exp\left(-\frac{r^2}{2}\right) r = g(r)\,h(t),$$

ahol

$$\begin{aligned} g(r) &\doteq r\exp\left(-\frac{r^2}{2}\right), \quad r>0, \\ h(t) &\doteq \frac{1}{2\pi}, \quad t\in(0,2\pi), \end{aligned}$$

vagyis a polárkoordináták is függetlenek, a φ szög a $(0,2\pi)$ intervallumban egyenletes eloszlású, a ρ sugár eloszlásának sűrűségfüggvénye pedig

$$g(r) = r\exp\left(-r^2/2\right).$$

A gondolatmenet megfordítható: ha a (ρ, φ) pár eloszlása éppen ilyen, akkor a (ξ, η) két független normális eloszlást definiál. Normális eloszlású változók szimulálása tehát visszavezethető egy egyenletes, és egy $r\exp\left(-r^2/2\right)$ sűrűségfüggvényű változó szimulálására. Ha a δ a $(0,1)$ intervallumon egyenletes eloszlású, akkor a $\rho \circeq \sqrt{-2\ln\delta}$ eloszlása

$$\mathbf{P}(\rho < r) = \mathbf{P}\left(\delta > \exp\left(-\frac{1}{2}r^2\right)\right) = 1 - \exp\left(-\frac{1}{2}r^2\right),$$

amely sűrűségfüggvénye éppen az $r\exp\left(-r^2/2\right)$.

Khi-négyzet eloszlás

Az n szabadságfokú χ_n^2 változót mint n darab független, standard normális eloszlású változó négyzetének összegét definiáljuk. A χ_n^2 eloszlását χ_n^2 eloszlásnak mondjuk. Ha $n = 1$, akkor, miként már láttuk, a $\chi_1^2 \cong N(0,1)^2$ sűrűségfüggvénye $\Gamma(1/2, 1/2)$

$$k_1(x) = \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{x}{2}\right), \quad x > 0.$$

Az általános esetben $\Gamma(n/2, n/2)$

$$k_n(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} \exp\left(-\frac{x}{2}\right), \quad x > 0.$$

amely a gamma eloszlás additív tulajdonsága miatt evidens.
A várható értéket és a szórás

$$\mathbf{E}\left(\chi_n^2\right) = n\mathbf{E}\left(\chi_1^2\right) = n, \quad \mathbf{D}^2\left(\chi_n^2\right) = n\mathbf{D}^2\left(\chi_1^2\right) = n(3-1) = 2n.$$

A χ_n^2 -ből gyökvonással

$$\begin{aligned} h_n(x) &= k_n\left(x^2\right)2x = \frac{\left(x^2\right)^{n/2-1}\exp\left(-x^2/2\right)}{2^{n/2}\Gamma(n/2)}2x = \\ &= \frac{1}{2^{n/2-1}\Gamma(n/2)}x^{n-1}\exp\left(-\frac{x^2}{2}\right). \end{aligned}$$

$$\mathbf{E}(\chi_n) = \int_0^\infty x h_n(x)\,dx = \frac{1}{2^{n/2-1}\Gamma(n/2)}\int_0^\infty x^n \exp\left(-\frac{x^2}{2}\right)dx.$$

A szokásos $t = x^2/2$ helyettesítéssel

$$\begin{aligned} \mathbf{E}(\chi_n) &= \frac{1}{2^{n/2-1}\Gamma(n/2)}\int_0^\infty 2^{n/2}t^{n/2}\exp(-t)\frac{1}{\sqrt{2t}}dt = \\ &= \frac{\sqrt{2}}{\Gamma(n/2)}\int_0^\infty t^{(n+1)/2-1}\exp(-t)\,dt = \sqrt{2}\frac{\Gamma((n+1)/2)}{\Gamma(n/2)}. \end{aligned}$$

$$\mathbf{D}^2(\chi_n) = \mathbf{E}\left(\chi_n^2\right) - \mathbf{E}(\chi_n)^2 = n - 2\frac{\Gamma((n+1)/2)^2}{\Gamma(n/2)^2}.$$

Student t

Legyen $\xi \cong N(0,1)$ és $t_n \doteq \xi / \chi_n$, ahol a χ_n-ről feltesszük, hogy független a ξ-től és értelemszerűen χ_n eloszlású. A t_n változó eloszlását n szabadságfokú Student-eloszlásnak mondjuk.

$$
\begin{aligned}
f_\xi(x) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \\
f_\eta(x) &= h_n(x) = \frac{1}{2^{n/2-1}\Gamma(n/2)} x^{n-1} \exp\left(-\frac{x^2}{2}\right), \quad x \geq 0
\end{aligned}
$$

A hányados sűrűségfüggvényének képlete szerint

$$
t_n(x) =
$$

$$
= \frac{1}{2^{n/2-1}\Gamma(n/2)} \frac{1}{\sqrt{2\pi}} \int_0^\infty y \exp\left(-\frac{(xy)^2}{2}\right) y^{n-1} \exp\left(-\frac{y^2}{2}\right) dy =
$$

$$
= \frac{1}{2^{n/2-1}\Gamma(n/2)} \frac{1}{\sqrt{2\pi}} \int_0^\infty y^n \exp\left(-y^2\left(1+x^2\right)/2\right) dy.
$$

A megszokott $u = y^2 \left(1 + x^2\right)/2$ helyettesítéssel

$$t_n\left(x\right) =$$

$$= \frac{1}{2^{n/2-1}\Gamma\left(n/2\right)} \frac{1}{\sqrt{2\pi}} \int_0^\infty \left(\sqrt{\frac{2u}{1+x^2}}\right)^n \exp\left(-u\right) \frac{dy}{du} du =$$

$$= \frac{1}{\sqrt{\pi}\Gamma\left(n/2\right)} \frac{1}{\left(1+x^2\right)^{(n+1)/2}} \Gamma\left(\frac{n}{2} - \frac{1}{2} + 1\right) =$$

$$= \frac{\Gamma\left(\left(n+1\right)/2\right)}{\sqrt{\pi}\Gamma\left(n/2\right)} \frac{1}{\left(1+x^2\right)^{(n+1)/2}}.$$

Cauchy eloszlás

Ha $n = 1$, akkor

$$t_1(x) = \frac{\Gamma(2/2)}{\sqrt{\pi}\Gamma(1/2)} \frac{1}{(1+x^2)} = \frac{1}{\sqrt{\pi}\sqrt{\pi}} \frac{1}{(1+x^2)} = \frac{1}{\pi(1+x^2)}.$$

Vagyis ha ξ és η független $N(0,1)$ eloszlásúak, akkor a $\xi / |\eta|$ hányados Cauchy-eloszlású. Ugyanakkor ha ξ és η normális eloszlásúak, akkor a ξ/η sűrűségfüggvénye

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{(xy)^2}{2}\right) \exp\left(-\frac{y^2}{2}\right) |y| \, dy =$$

$$= \frac{2}{2\pi} \int_0^{\infty} y \exp\left(-\frac{y^2(1+x^2)}{2}\right) dy =$$

$$= \frac{1}{\pi}\left[-\frac{1}{1+x^2} \exp\left(-\frac{y^2(1+x^2)}{2}\right)\right]_0^{\infty} = \frac{1}{\pi} \frac{1}{1+x^2},$$

amely szintén Cauchy eloszlású.

Mivel a sűrűségfüggvény páros, és így a várható érték, ha létezik, csak nulla lehet. Ha azonban $n = 1$, akkor nincs várható érték. Ha $n \geq 2$, akkor

$$\mathbf{D}^2(t_n) =$$
$$= \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)} \int_{\mathbb{R}} \frac{x^2}{\left(1+x^2\right)^{(n+1)/2}} dx =$$
$$= \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)} 2 \int_0^\infty \frac{x^2}{\left(1+x^2\right)^{(n+1)/2}} dx.$$

Student t eloszlás

Ha $t = x^2 / \left(1 + x^2\right)$, vagyis ha $x = \sqrt{t / (1-t)}$, akkor

$$\mathbf{D}^2 (t_n) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)} 2 \int_0^1 t\left(1-t\right)^{(n-1)/2} \frac{dx}{dt} dt =$$

$$= \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)} \int_0^1 t\left(1-t\right)^{(n-1)/2} \frac{\sqrt{1-t}}{\sqrt{t}} \frac{1}{\left(1-t\right)^2} dt =$$

$$= \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)} \int_0^1 t^{1/2}\left(1-t\right)^{n/2-2} dt =$$

$$= \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)} B\left(\frac{3}{2}, \frac{n}{2} - 1\right).$$

Student t eloszlás

A levezetésből világos, hogy amennyiben $n/2-1 \leq 0$, vagyis $n<3$, akkor a szórás végtelen. Ha $n \geq 3$, akkor

$$\begin{aligned}
\mathbf{D}^2\left(t_n\right) &= \frac{1}{\sqrt{\pi}}\frac{\Gamma\left(\left(n+1\right)/2\right)}{\Gamma\left(n/2\right)}\frac{\Gamma\left(1/2+1\right)\Gamma\left(n/2-1\right)}{\Gamma\left(3/2+\left(n-2\right)/2\right)} = \\
&= \frac{\frac{1}{2}\sqrt{\pi}\Gamma\left(n/2-1\right)}{\sqrt{\pi}\Gamma\left(n/2\right)} = \\
&= \frac{1}{2}\frac{\Gamma\left(n/2-1\right)}{\left(n/2-1\right)\Gamma\left(n/2-1\right)} = \frac{1}{n-2}.
\end{aligned}$$

Student t eloszlás

Esetenként Student-eloszláson a

$$\widetilde{t}_n = \sqrt{n} t_n = \sqrt{n} \frac{\xi}{\chi_n} = \frac{\xi}{\sqrt{\left(\xi_1^2 + \ldots + \xi_n^2\right) / n}}$$

változó eloszlását szokás érteni. A t_n eloszlásából a $\widetilde{t}_n$ eloszlása

$$\widetilde{F}(x) = \mathbf{P}\left(\widetilde{t}_n < x\right) = \mathbf{P}\left(\sqrt{n} t_n < x\right) = F\left(\frac{x}{\sqrt{n}}\right),$$

amiből a sűrűségfüggvény

$$\frac{\Gamma\left((n+1)/2\right)}{\sqrt{\pi n} \Gamma\left(n/2\right)} \frac{1}{\left(1 + x^2/n\right)^{(n+1)/2}}.$$

A már belátottak alapján világos, hogy ha $n \geq 2$, akkor a várható érték nulla, ha pedig $n \geq 3$, akkor a variancia

$$\mathbf{D}^2\left(\widetilde{t}_n\right) = \frac{n}{n-2}.$$

Student t próba

Lemma

Legyenek $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ *független* $N(0,1)$ *változók. Legyen* $\mathbf{C}$ *ortogonális mátrix. Ha* $\boldsymbol{\eta} = \mathbf{C}\boldsymbol{\xi}$ *akkor az* $\eta_1, \ldots, \eta_n$ *változók is független* $N(0,1)$ *változók és*

$$\sum_{k=1}^{n} \xi_k^2 = \sum_{k=1}^{n} \eta_k^2.$$

Ha a változók függetlenek, akkor a többdimenziós eloszlás is normális. Többdimenziós normális eloszlások lineáris kombinációja is normális. A várható érték triviálisan nulla. Normális eloszlású változók esetén a függetlenség következik a korrelálatlanságból.

$$\begin{aligned}
\mathbf{E}\left(\eta_k \eta_j\right) &= \mathbf{E}\left(\sum_i c_{ki}\xi_i \sum_l c_{jl}\xi_l\right) = \\
&= \sum_i \sum_l c_{ki} c_{jl} \mathbf{E}\left(\xi_i \xi_l\right) = \sum_i c_{ki} c_{ji} = \begin{cases} 0 & \text{ha} \quad k \neq j \\ 1 & \text{ha} \quad k = j \end{cases}.
\end{aligned}$$

Student t próba

Hasonlóan

$$\begin{aligned}\sum_{k=1}^{n}\eta_k^2 &= \sum_{k=1}^{n}\left(\sum_{i=1}^{n}c_{ki}\xi_i\right)^2 = \sum_{k=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ki}c_{kj}\xi_i\xi_j = \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i\xi_j\sum_{k=1}^{n}c_{ki}c_{kj} = \sum_{i=1}^{n}\xi_i^2.\end{aligned}$$

Student t próba

Tekintsük a $\overline{\xi} = \frac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ átlagot és az

$$\begin{aligned} s^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\xi_i^2 + \overline{\xi}^2 - 2\xi_i\overline{\xi}\right) = \\ &= \frac{1}{n}\sum_{i=1}^{n}\xi_i^2 + \overline{\xi}^2 - 2\overline{\xi}^2 = \frac{1}{n}\sum_{i=1}^{n}\xi_i^2 - \overline{\xi}^2 \end{aligned}$$

empririkus szórásnégyzetet. Legyen **C** olyan ortogonális mátrix, amely első sora csupa $1/\sqrt{n}$. Ekkor

$$\begin{aligned} \overline{\xi} &= \frac{1}{\sqrt{n}}\eta_1, s^2 = \frac{1}{n}\sum_{i=1}^{n}\xi_i^2 - \overline{\xi}^2 = \frac{1}{n}\sum_{k=1}^{n}\eta_k^2 - \frac{1}{n}\eta_1^2 = \\ &= \frac{1}{n}\sum_{k=2}^{n}\eta_k^2, \end{aligned}$$

vagyis a $\overline{\xi}$ és az s^2 függetlenek, a $\overline{\xi}$ eloszlása $N(0,1)/\sqrt{n}$ és az ns^2 tőle független χ^2_{n-1}.

Student t próba

Így a

$$t_{n-1} = \frac{\sqrt{n}\overline{\xi}}{\sqrt{ns^2}} = \frac{\overline{\xi}}{s} = \frac{\overline{\xi}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2}}.$$

$n-1$ szabadságfokú Student t eloszlás. Illetve

$$\widetilde{t}_{n-1} = \sqrt{n-1}\,t_{n-1} = \frac{\overline{\xi}}{\sqrt{\frac{1}{(n-1)n}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2}}.$$

Ez utóbbi szokásos jelölése

$$\widetilde{t}_{n-1} = \frac{\overline{\xi}}{S/\sqrt{n}}$$

ahol $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2}$.

F-eloszlás

Az m, n szabadságfokú Fisher-féle F eloszláson az

$$F(m,n) \circeq \frac{1/m\sum_{i=1}^{m}\xi_i^2}{1/n\sum_{i=1}^{n}\eta_i^2}$$

változó eloszlását értjük, ahol a ξ_i és az η_j független $N(0,1)$ eloszlású változók. Legyen $\kappa = (m/n)\,F(m,n)$. A κ két független χ^2 eloszlás hányadosa.

Theorem

Ha $\xi \cong \Gamma(a,\lambda)$ és $\eta \cong \Gamma(b,\lambda)$ valamint a ξ és az η függetlenek, akkor

$$\frac{\xi}{\eta} \cong \widetilde{B}(a,b), \quad \frac{\xi}{\xi+\eta} \cong B(a,b).$$

F-eloszlás

Definition

Ha α és β pozitív paraméterek, akkor a

$$g(x) \doteq \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} \frac{1}{(1+x)^{\alpha+\beta}}, \quad x > 0$$

sűrűségfüggvénnyel rendelkező eloszlást általánosított, vagy másodfajú béta eloszlásnak nevezzük. Az eloszlást $\widetilde{B}(\alpha, \beta)$-val fogjuk jelölni. Béta eloszláson a

$$h(x) \doteq \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 < x < 1$$

sűrűségfüggvénnyel rendelkező eloszlást értjük.

F-eloszlás

Bizonyítás: A hányados valószínűségi változó sűrűségfüggvényének képletét felírva, és kihasználva a két változó nem negativitását, ha $u > 0$, akkor

$$\int_0^\infty \frac{\lambda^a}{\Gamma(a)} (uy)^{a-1} \exp(-\lambda uy) \frac{\lambda^b}{\Gamma(b)} y^{b-1} \exp(-\lambda y)\, y dy =$$

$$\begin{aligned}
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} u^{a-1} \int_0^\infty y^{a+b-1} \exp(-\lambda y(u+1))\, dy = \\
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} u^{a-1} \int_0^\infty \left(\frac{t}{\lambda(1+u)}\right)^{a+b-1} \exp(-t) \frac{1}{\lambda(1+u)} dt = \\
&= \frac{1}{\Gamma(a)\Gamma(b)} u^{a-1} \frac{1}{(1+u)^{a+b}} \int_0^\infty t^{a+b-1} \exp(-t)\, dt = \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1} \frac{1}{(1+u)^{a+b}},
\end{aligned}$$

F-eloszlás

Mivel

$$\frac{\xi}{\eta+\xi}=\frac{\xi/\eta}{1+\xi/\eta}$$

elegendő a ξ/η változón $\varphi(x)=x/(1+x)$ transzformációt végezni.

$$\varphi^{-1}(x)=x/(1-x),0<x<1.$$

A transzformált sűrűségfüggvény

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left(\frac{x}{1-x}\right)^{a-1}\frac{1}{\left(1+\frac{x}{1-x}\right)^{a+b}}\frac{1}{(1-x)^2}=$$

$$=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left(\frac{x}{1-x}\right)^{a-1}(1-x)^{a+b}\frac{1}{(1-x)^2}=$$

$$=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}.$$

Többdimenziós normális eloszlás

A többdimenziós normális eloszlás definíciójánál a probléma abból származik, hogy egy egységes tárgyalást akarunk adni az elfajult és a nem elfajult esetre. Az m-dimenziós elfajult esetben az eloszlás az $\mathbb{R}^m$ egy valódi alterére koncentrálódik, ezért nincs sűrűségfüggvénye. A többdimenziós normális eloszlás definícióját visszavezetjük az egydimenziós esetre.

Definition

Az $\mathbb{R}^m$ értékű $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$ változót normálisnak mondjuk, ha tetszőleges $\mathbf{t} \in \mathbb{R}^m$ vektorra a $(\mathbf{t}, \boldsymbol{\xi}) \overset{\circ}{=} \sum_{i=1}^{m} t_i \xi_i$ skaláris szorzat eloszlása normális.

A definíció szerint tetszőleges $\mathbf{t}$ vektorra a $(\mathbf{t}, \boldsymbol{\xi})$ eloszlását egyértelműen meghatározza az $m_{\mathbf{t}}$ várható értéke és $\sigma_{\mathbf{t}}$ szórása. Könnyen látható, hogy $m_{\mathbf{t}} = (\mathbf{m}, \mathbf{t})$, ahol $\mathbf{m}$ a $\boldsymbol{\xi}$ várható értéke, valamint

$$\sigma_{\mathbf{t}}^2 = \mathbf{E}\left((\mathbf{t}, \boldsymbol{\xi})^2\right) - (\mathbf{m}, \mathbf{t})^2 = \left(\mathbf{t}, \mathbf{E}\left(\boldsymbol{\xi}\boldsymbol{\xi}^T\right)\mathbf{t}\right) - (\mathbf{m}, \mathbf{t})^2 = (\mathbf{t}, \mathbf{S}\mathbf{t}),$$

ahol az $\mathbf{S}$ mátrix s_{ij} eleme éppen $\mathbf{E}\left((\xi_i - m_i)\left(\xi_j - m_j\right)\right)$, vagyis a ξ_i és ξ_j változók kovarianciája. Az $\mathbf{S}$ evidens módon pozitív szemidefinít.

Többdimenziós normális eloszlás

Az egydimenziós normális eloszlás karakterisztikus függvényének képlete alapján a $(\mathbf{t}, \boldsymbol{\xi})$ szorzat eloszlásának karakterisztikus függvénye

$$\varphi_{\mathbf{t}}(s) = \exp\left(im_{\mathbf{t}}s - \sigma_{\mathbf{t}}^2 \frac{s^2}{2}\right) = \exp\left(i(\mathbf{t}, \mathbf{m})\, s - (\mathbf{t}, \mathbf{St})\frac{s^2}{2}\right),$$

amely az $s = 1$ helyen éppen a $\boldsymbol{\xi}$ karakterisztikus függvénye a $\mathbf{t}$ helyen, amiből

$$\varphi(\mathbf{t}) = \exp\left(i(\mathbf{t}, \mathbf{m}) - \frac{1}{2}(\mathbf{t}, \mathbf{St})\right).$$

Többdimenziós normális eloszlás

A többdimeziós normális eloszlás egy másik szokásos bevezetése szerint az eloszlást a karakterisztikus függvényén keresztül érdemes definiálni. A többdimenziós normális eloszlással kapcsolatos bizonyításokban legtöbbször az eloszlás helyett a karakterisztikus függvényét vizsgáljuk, így ez a megközelítés igen célratörő.

Definition

A $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$ vektor eloszlását normálisnak mondjuk, ha a karakterisztikus függvénye

$$\varphi(\mathbf{t}) = \exp\left(i(\mathbf{t}, \mathbf{m}) - \frac{1}{2}(\mathbf{t}, \mathbf{St})\right)$$

alakú, ahol az $\mathbf{S}$ $(m \times m)$-es pozitív szemidefinít mátrix, az $\mathbf{m}$ m-dimenziós vektor. Az $(\mathbf{m}, \mathbf{S})$ paraméterekhez tartozó többdimenziós normális eloszlást $N(\mathbf{m}, \mathbf{S})$ módon jelöljük

.

Az alábbiak során többször fogunk hivatkozni a következő lineáris algebrai állításra:

Theorem (Spektrálfelbontási tétel)

Tetszőleges **S** *szimmetrikus, valós értékű mátrixhoz létezik az* **S** *sajátvektoraiból alkotott olyan ortonomált bázis, amelyben az* **S** *diagonális alakra hozható, vagyis az* **S** *mátrixhoz található olyan* **C** *ortonormált mátrix, hogy*

$$\mathbf{D} = \mathbf{C}^T \mathbf{S} \mathbf{C} = \mathbf{C}^{-1} \mathbf{S} \mathbf{C},$$

ahol a **D** *diagonális, és a diagonálisban levő elemek éppen az* **S** *sajátértékei.*

Többdimenziós normális eloszlás

.

Bizonyítás: Miként ismert az algebra alaptétele miatt tetszőleges lineáris leképezés rendelkezik komplex sajátértékekkel, vagyis tetszőleges $\mathbf{S}$ mátrix esetén van olyan $\lambda \in \mathbb{C}$ szám és $\mathbf{x} \neq \mathbf{0}$ vektor, mindkettő esetlegesen komplex, amelyekre $\mathbf{Sx} = \boldsymbol{\lambda}\mathbf{x}$. Ha λ komplex lenne, akkor a konjugáltja $\overline{\lambda}$ szintén sajátérték lenne, ugyanis az $\mathbf{Sx} = \lambda_1\mathbf{x}$ egyenletet konjugálva, felhasználva, hogy az $\mathbf{S}$ valós

$$\mathbf{S}\overline{\mathbf{x}} = \overline{\mathbf{S}}\overline{\mathbf{x}} = \overline{\mathbf{Sx}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}.$$

Ha $\mathbf{x}$ sajátvektor, akkor

$$\begin{aligned} \lambda\left(\mathbf{x},\mathbf{x}\right) &= \left(\mathbf{x},\boldsymbol{\lambda}\mathbf{x}\right) = \left(\mathbf{x},\mathbf{Sx}\right) = \left(\overline{\mathbf{S}}^{T}\mathbf{x},\mathbf{x}\right) = \\ &= \left(\mathbf{Sx},\mathbf{x}\right) = \left(\lambda\mathbf{x},\mathbf{x}\right) = \overline{\lambda}\left(\mathbf{x},\mathbf{x}\right), \end{aligned}$$

amiből az $(\mathbf{x},\mathbf{x}) \neq 0$ miatt $\lambda = \overline{\lambda}$, vagyis a λ valós. Ebből következően nem csak a sajátértékek, hanem a sajátvektorok is valósak.

.

Ha λ_1 és λ_2 az $\mathbf{S}$ szimmetrikus, valós értékű mátrix két különböző sajátértéke és $\lambda_2 \neq 0$ és $\mathbf{x}_1$, illetve $\mathbf{x}_2$ a megfelelő sajátvektorok, akkor

$$\begin{aligned}(\mathbf{x}_1, \mathbf{x}_2) &= \left(\mathbf{x}_1, \frac{1}{\lambda_2}\mathbf{S}\mathbf{x}_2\right) = \left(\mathbf{S}^T\mathbf{x}_1, \frac{1}{\lambda_2}\mathbf{x}_2\right) = \\ &= \frac{1}{\lambda_2}(\mathbf{S}\mathbf{x}_1, \mathbf{x}_2) = \left(\lambda_1\mathbf{x}_1, \frac{1}{\lambda_2}\mathbf{x}_2\right) = \\ &= \frac{\lambda_1}{\lambda_2}(\mathbf{x}_1, \mathbf{x}_2),\end{aligned}$$

ami csak akkor lehetséges, ha $(\mathbf{x}_1, \mathbf{x}_2) = 0$..

.

Ha $\mathbf{x}_1 \neq 0$ egy sajátvektor, akkor $\mathbf{x}_1 / \left\|\mathbf{x}_1\right\|$ egységnyi hosszú sajátvektor. Tegyük fel, hogy $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ ortonormált sajátvektorai az $\mathbf{S}$-nak. Legyen $L \stackrel{\circ}{=} \operatorname{lin}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ a kifeszített altér. Ha az L dimenziója nem azonos az $\mathbf{S}$ dimenziójával, akkor az $L^{\perp} \stackrel{\circ}{=} \{\mathbf{x} : (\mathbf{x}, \mathbf{x}_i) = 0, i = 1, 2, \ldots, k\}$ altér nem üres. Ha $\mathbf{x} \in L^{\perp}$, akkor

$$(\mathbf{Sx}, \mathbf{x}_i) = \left(\mathbf{x}, \mathbf{S}^T \mathbf{x}_i\right) = (\mathbf{x}, \mathbf{Sx}_i) = \lambda_i (\mathbf{x}, \mathbf{x}_i) = 0,$$

vagyis az $\mathbf{S}$ az $L^{\perp}$ alteret önmagára képezi. Így az $\mathbf{S}$-nak van sajátvektora az $L^{\perp}$ altérben. Ha $\mathbf{x}_{k+1}$ a megfelelő normalizált sajátvektor, akkor $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k+1}$ egy $k+1$ elemből álló ortonormált sajátvektor rendszer.

.

Legyen az $\mathbf{S}$ dimenziója n. Az elmondottak alapján léteznek az $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ ortonormált sajátvektorok. Ha $\mathbf{C}$ oszlopai éppen az $(\mathbf{x}_k)_{k=1}^n$ sajátvektorok és a $\mathbf{D}$ diagonális mátrix átlójában éppen a sajátértékek vannak, akkor $\mathbf{SC} = \mathbf{CD}$, amit $\mathbf{C}^T$-vel megszorozva

$$\mathbf{C}^T\mathbf{SC} = \mathbf{C}^T(\mathbf{CD}) = \left(\mathbf{C}^T\mathbf{C}\right)\mathbf{D} = \mathbf{D},$$

ami éppen a kívánt összefüggés.

.

Corollary

Ha $\mathbf{S}$ *olyan* $(m \times m)$*-es szimmetrikus, pozitív szemidefinit mátrix, amely rangja* r*, akkor létezik olyan* $\mathbf{A}$ $(m \times r)$*-es mátrix, amelyre*

$$\mathbf{S} = \mathbf{A}\mathbf{A}^T.$$

Az $\mathbf{A}$ *mátrix oszlopai az* $\mathbf{S}$ *sajátvektorai és az* $\mathbf{A}$ *minden* $\mathbf{a}_j$ *oszlopára*

$$\|\mathbf{a}_j\| = \sqrt{\lambda_j},$$

ahol λ_j *az* $\mathbf{S}$ *valamelyik nullától különböző sajátértéke.*

.

Bizonyítás: A spektrálfelbontási tétel alapján az $\mathbf{S}$ sajátvektoraiból készíthető ortonormált bázis. Ha a bázisvektorokból egy $\mathbf{C}$ mátrixot képezünk, akkor a sajátvektor definíciójából $\mathbf{SC} = \mathbf{CD}$, ahol $\mathbf{D}$ a sajátértékekből álló diagonális mátrix. Felhasználva a sajátvektorok ortonormalitását, $\mathbf{S} = \mathbf{CDC}^T$. Mivel az $\mathbf{S}$ pozitív szemidefinit, ezért a $\mathbf{D}$ átlójában szereplő sajátértékek nem negatívak, tehát képezhetjük a négyzetgyökeikből álló $\mathbf{D}^{1/2}$ mátrixot. Evidens módon a $\mathbf{B} \triangleq \mathbf{CD}^{1/2}$ mátrixra

$$\begin{aligned}
\mathbf{BB}^T &\triangleq \mathbf{CD}^{1/2}\left(\mathbf{CD}^{1/2}\right)^T = \mathbf{CD}^{1/2}\left(\mathbf{D}^{1/2}\right)^T\mathbf{C}^T = \\
&= \mathbf{CD}^{1/2}\mathbf{D}^{1/2}\mathbf{C}^T = \mathbf{CDC}^T = \mathbf{S}.
\end{aligned}$$

.

Mivel az $\mathbf{S}$ rangja r, ezért a $\mathbf{D}$ átlójában pontosan r pozitív elem található. Alkalmas indexelés mellett $\mathbf{B} \overset{\circ}{=} \begin{pmatrix} \mathbf{A} & \mathbf{0} \end{pmatrix}$, ahol az $\mathbf{A}$ dimenziója $(m \times r)$. Evidens módon

$$\mathbf{S} = \mathbf{B}\mathbf{B}^T = \begin{pmatrix} \mathbf{A} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{A}^T \\ \mathbf{0} \end{pmatrix} = \mathbf{A}\mathbf{A}^T.$$

Többdimenziós normális eloszlás

.

Theorem

Teljesülnek az alábbi állítások:

- *Ha* $\mathbf{m}$ *tetszőleges* $\mathbb{R}^m$*-beli vektor, és* $\mathbf{S}$ $(m \times m)$*-es pozitív szemidefinit mátrix, akkor az* $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ *téren létezik olyan eloszlás, amelynek karakterisztikus függvénye éppen a megadott.*
- *Az* $N(\mathbf{m}, \mathbf{S})$ *eloszlás várható érték vektora* $\mathbf{m}$, *kovariancia mátrixa* $\mathbf{S}$.
- *Ha* $\xi \cong N(\mathbf{m}, \mathbf{S})$ *és* $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ *lineáris leképezés, akkor*

$$\mathbf{T}\xi \cong N\left(\mathbf{Tm}, \mathbf{TST}^T\right).$$

Többdimenziós normális eloszlás

.

Theorem

- *Ha* $\xi \cong N(\mathbf{m}, \mathbf{S})$, *az* $\mathbf{S}$ *mátix diagonális és a diagonális elemek rendre* σ_i^2, $(i = 1, \ldots, m)$, *akkor az egyes* ξ_i *változók* $N(m_i, \sigma_i)$ *egydimenziós normális eloszlásúak, ahol feltételezzük, hogy definíció szerint a nulla szórású egydimenziós normális eloszlású valószínűségi változók majdnem mindenhol konstansok. Általában tetszőleges* $\mathbf{t}$ *vektorra*

$$(\mathbf{t}, \xi) \cong N\left((\mathbf{t}, \mathbf{m}), (\mathbf{t}, \mathbf{St})^{1/2}\right).$$

Többdimenziós normális eloszlás

.

Theorem

- *Ha* $\boldsymbol{\xi} \cong N(\mathbf{m}, \mathbf{S})$ *és az* $\mathbf{S}$ *kovariancia mátrix rangja* r, *akkor létezik olyan* $\mathbf{A}$ $(m \times r)$ *mátrix, és* $\boldsymbol{\eta} \cong N(\mathbf{0}, \mathbf{E}_r)$, *ahol* $\mathbf{E}_r$ *az* r*-dimenziós egységmátrix, hogy* $\boldsymbol{\xi} = \mathbf{A}\boldsymbol{\eta} + \mathbf{m}$. *Az* $\mathbf{A}$ *mátrix oszlopai az* $\mathbf{S}$ *olyan sajátvektorai, amelyekhez az* $\mathbf{S}$ *nem nulla sajátértékei tartoznak. Az* $\mathbf{A}$ *mátrix oszlopainak hossza éppen a megfelelő pozitív sajátérték négyzetgyöke.*
- *Ha az* $\mathbf{S}$ *kovariancia mátrixa szigorúan pozitív definit, vagyis ha az* $\mathbf{S}$ *mátrix invertálható, akkor az* $N(\mathbf{m}, \mathbf{S})$ *eloszlás abszolút folytonos a* λ_m *Lebesgue-mértékre nézve, vagyis az eloszlásnak van sűrűségfüggvénye. A sűrűségfüggvény képlete:*

$$f(\mathbf{x}) = \frac{1}{\left(\sqrt{2\pi}\right)^n} \frac{1}{\sqrt{\det(\mathbf{S})}} \exp\left(-\frac{1}{2}\left(\mathbf{x} - \mathbf{m}, \mathbf{S}^{-1}(\mathbf{x} - \mathbf{m})\right)\right).$$

.

Theorem

- *Ha* $\boldsymbol{\xi} \cong N(\mathbf{m}, \mathbf{S})$, $\mathbf{S} \stackrel{\circ}{=} (s_{ij})$ *és* $s_{ij} = 0$, *ha* $i \neq j$, *akkor a* ξ_i *és* ξ_j *változók függetlenek.*
- *Ha valamely* $\mathbf{W}$ *mátrixra* $\mathbf{S} = \mathbf{W}\mathbf{W}^T$, *és a* $\boldsymbol{\xi}$ *várható értéke nulla, akkor a* $\boldsymbol{\xi}$ 1 *valószínűséggel eleme a* $\mathbf{W}$ *oszlopvektorai által kifeszített* $H \stackrel{\circ}{=} \mathrm{lin}(\mathbf{W})$ *altérnek. Ha* $\mathbf{M}(\boldsymbol{\xi}) = \mathbf{m}$, *akkor* $\mathbf{P}(\boldsymbol{\xi} \in \mathbf{m} + H) = 1$.

Többdimenziós normális eloszlás

.

Bizonyítás: Ha az $\mathbf{S}$ rangja r, akkor létezik $(m \times r)$-es $\mathbf{A}$ mátrix, hogy $\mathbf{S} = \mathbf{A}\mathbf{A}^T$. Tetszőleges η r-dimenziós változóra, ha $\xi \overset{\circ}{=} \mathbf{A}\eta + \mathbf{m}$, akkor a ξ karakterisztikus függvénye

$$\varphi_\xi(\mathbf{t}) = \mathbf{E}\left(\exp\left(i\left(\mathbf{t}, \xi\right)\right)\right) = \mathbf{E}\left(\exp\left(i\left(\mathbf{t}, \mathbf{A}\eta\right) + i\left(\mathbf{t}, \mathbf{m}\right)\right)\right) =$$

$$= \mathbf{E}\left(\exp\left(i\left(\mathbf{A}^T\mathbf{t}, \eta\right) + i\left(\mathbf{t}, \mathbf{m}\right)\right)\right) = \exp\left(i\left(\mathbf{t}, \mathbf{m}\right)\right)\varphi_\eta\left(\mathbf{A}^T\mathbf{t}\right).$$

.

Ha η r darab független, standard normális eloszlású változóból álló vektor, akkor

$$\varphi_{\eta}\left(\mathbf{s}\right)=\prod_{i=1}^{r}\varphi_{i}\left(s_{i}\right)=\prod_{i=1}^{r}\exp\left(-\frac{1}{2}s_{i}^{2}\right)=\exp\left(-\frac{1}{2}\left(\mathbf{s},\mathbf{s}\right)\right),$$

amiből

$$\begin{aligned}\varphi_{\xi}\left(\mathbf{t}\right) &= \exp\left(i\left(\mathbf{t},\mathbf{m}\right)-\frac{1}{2}\left(\mathbf{A}^{T}\mathbf{t},\mathbf{A}^{T}\mathbf{t}\right)\right)= \\ &= \exp\left(i\left(\mathbf{t},\mathbf{m}\right)-\frac{1}{2}\left(\mathbf{t},\mathbf{A}\mathbf{A}^{T}\mathbf{t}\right)\right)= \\ &= \exp\left(i\left(\mathbf{t},\mathbf{m}\right)-\frac{1}{2}\left(\mathbf{t},\mathbf{S}\mathbf{t}\right)\right),\end{aligned}$$

ami éppen az első állítás.

.

Elemi számolással látható, hogy

$$\begin{aligned}\mathbf{E}(\xi) &= \mathbf{E}(\mathbf{A}\eta+\mathbf{m}) = \mathbf{A}\cdot\mathbf{E}(\eta)+\mathbf{m} = \mathbf{m},\\ \mathbf{E}\left((\xi-\mathbf{m})(\xi-\mathbf{m})^{T}\right) &= \mathbf{E}\left((\mathbf{A}\eta)(\mathbf{A}\eta)^{T}\right) = \mathbf{E}\left(\mathbf{A}\eta\eta^{T}\mathbf{A}^{T}\right) =\\ &= \mathbf{A}\cdot\mathbf{E}\left(\eta\eta^{T}\right)\cdot\mathbf{A}^{T} = \mathbf{A}\mathbf{E}_{r}\mathbf{A}^{T} = \mathbf{S},\end{aligned}$$

vagyis teljesül a második állítás is.

Többdimenziós normális eloszlás

.

$$\begin{aligned}
\varphi_{\eta}(\mathbf{u}) &= \mathbf{E}\left(\exp\left(i\left(\mathbf{u},\mathbf{T}\boldsymbol{\xi}\right)\right)\right) = \mathbf{E}\left(\exp\left(i\left(\mathbf{T}^{T}\mathbf{u},\boldsymbol{\xi}\right)\right)\right) = \\
&= \varphi_{\xi}\left(\mathbf{T}^{T}\mathbf{u}\right) = \exp\left(\left(\mathbf{T}^{T}\mathbf{u},\mathbf{m}\right) + \frac{1}{2}\left(\mathbf{T}^{T}\mathbf{u},\mathbf{STu}\right)\right) = \\
&= \exp\left(\left(\mathbf{u},\mathbf{Tm}\right) + \frac{1}{2}\left(\mathbf{u},\mathbf{TST}^{T}\mathbf{u}\right)\right).
\end{aligned}$$

Többdimenziós normális eloszlás

.

Vegyük a $\mathbf{t} \mapsto (\mathbf{t}, \boldsymbol{\xi})$ lineáris leképezést. Az előző alapján az $\eta = (\mathbf{t}, \boldsymbol{\xi})$ eloszlása

$$N\left((\mathbf{t}, \mathbf{m}), (\mathbf{t}, \mathbf{St})\right),$$

amely karakterisztikus függvénye

$$\psi(s) = \exp\left(s(\mathbf{t}, \mathbf{m}) - \frac{1}{2}s^2(\mathbf{t}, \mathbf{St})\right).$$

Ha $\mathbf{t} = \mathbf{e}_j$, akkor

$$\psi(s) = \exp\left(m_j - \frac{1}{2}s^2\sigma_j^2\right).$$

Tehát a $\xi_j \cong N(m_j, \sigma_j)$.

.
Feltehetjük, hogy a ξ várható értéke nulla. Legyen $\mathbf{D} \stackrel{\circ}{=} \mathbf{C}^T\mathbf{SC}$ az $\mathbf{S}$ spektrálfelbontása. Ha

$$\zeta \stackrel{\circ}{=} \mathbf{C}^T\xi = \mathbf{C}^{-1}\xi,$$

akkor a harmadik állítás alapján a ζ kovariancia mátrixa

$$\mathbf{C}^T\mathbf{SC}^{TT} = \mathbf{C}^T\mathbf{SC} = \mathbf{D},$$

vagyis diagonális. Az $\mathbf{S}$ rangjára tett feltétel szerint feltehető, hogy a $\mathbf{D}$ diagonálisában levő első r elem pozitív, a többi pedig nulla. Mivel a ζ várható érték vektora nulla, ezért $\zeta \stackrel{\circ}{=} \begin{pmatrix} \eta_1 \\ \mathbf{0} \end{pmatrix}$, ahol az η_1 r-dimenziós normális eloszlású vektor. Ebből

$$\xi = \mathbf{C}\zeta = \begin{pmatrix} \mathbf{U} & \mathbf{V} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \mathbf{0} \end{pmatrix} = \mathbf{U}\eta_1.$$

Az $\mathbf{U}$ oszlopait az η_1 szórásával, vagyis $\sqrt{d_{ii}}$-vel, megszorozva az η_1 standardizálható, vagyis az ötödik állítás teljesül.

Többdimenziós normális eloszlás

.

Ha $\mathbf{S}$ invertálható, és $\mathbf{S} = \mathbf{A}\mathbf{A}^T$, akkor az $\mathbf{A}$ is invertálható. Ha $\boldsymbol{\xi} = \mathbf{A}\boldsymbol{\eta} + \mathbf{m}$, ahol $\boldsymbol{\eta} \cong N(\mathbf{0}, \mathbf{E})$, akkor $\boldsymbol{\xi} \cong N(\mathbf{m}, \mathbf{S})$. Az integráltranszformációs-tétel alapján minden B Borel-halmazra

$$
\begin{aligned}
\mathbf{P}(\xi \in B) &= \mathbf{P}(\mathbf{A}\boldsymbol{\eta} + \mathbf{m} \in B) = \mathbf{P}\left(\boldsymbol{\eta} \in \mathbf{A}^{-1}(B - \mathbf{m})\right) = \\
&= \int_{\mathbf{A}^{-1}(B-\mathbf{m})} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} s_i^2\right) d\mathbf{s} = \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^n} \int_{\mathbf{A}^{-1}(B-\mathbf{m})} \exp\left(-\frac{(\mathbf{s}, \mathbf{s})}{2}\right) d\mathbf{s} = \\
&= \frac{\left|\det \mathbf{A}^{-1}\right|}{\left(\sqrt{2\pi}\right)^n} \int_B \exp\left(-\frac{\left(\mathbf{A}^{-1}(\mathbf{s} - \mathbf{m}), \mathbf{A}^{-1}(\mathbf{s} - \mathbf{m})\right)}{2}\right) d\mathbf{s} =
\end{aligned}
$$

.

$$= \frac{\left|\det \mathbf{A}^{-1}\right|}{\left(\sqrt{2\pi}\right)^n} \int_B \exp\left(-\frac{\left(\left(\mathbf{s}-\mathbf{m}\right), \left(\mathbf{A}^{-1}\right)^T \mathbf{A}^{-1}\left(\mathbf{s}-\mathbf{m}\right)\right)}{2}\right) d\mathbf{s} =$$

$$= \frac{\left|\det \mathbf{A}^{-1}\right|}{\left(\sqrt{2\pi}\right)^n} \int_B \exp\left(-\frac{\left(\left(\mathbf{s}-\mathbf{m}\right), \left(\mathbf{A}\mathbf{A}^T\right)^{-1}\left(\mathbf{s}-\mathbf{m}\right)\right)}{2}\right) d\mathbf{s} =$$

$$= \frac{\left|\det \mathbf{A}^{-1}\right|}{\left(\sqrt{2\pi}\right)^n} \int_B \exp\left(-\frac{1}{2}\left(\left(\mathbf{s}-\mathbf{m}\right), \mathbf{S}^{-1}\left(\mathbf{s}-\mathbf{m}\right)\right)\right) d\mathbf{s}.$$

Mivel $\mathbf{S} = \mathbf{A}\mathbf{A}^T$, ezért $\left|\det \mathbf{A}^{-1}\right| = 1/\sqrt{\left|\det \mathbf{S}\right|}$, következésképpen a ξ sűrűségfüggvénye éppen az állításban szereplő alakkal rendelkezik.

Többdimenziós normális eloszlás

.

A (ξ_i, ξ_j) pár eloszlásának karakterisztikus függvénye az eredetiből úgy kapható, hogy a többi változóhoz tartozó t_k $(k \neq i, j)$ értékét nullának válasszuk. Ha $i \neq j$, és $s_{ij} = 0$, akkor a (ξ_i, ξ_j) eloszlásának karakterisztikus függvényében szereplő kovariancia mátrix diagonális. Ha a két eloszlás egyike sem elfajuló, akkor a kétváltozós eloszlásnak van sűrűségfüggvénye, és ennek alakjából a függetlenség már következik, hiszen a két dimenziós eloszlás sűrűségfüggvénye az egyes peremeloszlások sűrűségfüggvényének szorzata. Ha valamelyik eloszlás elfajult, akkor az állítás következik abból, hogy minden konstans valószínűségi változó minden más változótól független.

Többdimenziós normális eloszlás

.

Legyen η független standard normális eloszlású változókból álló oszlopvektor. Az elmondottak alapján tetszőleges $\mathbf{W}$ mátrixra a $\zeta \doteq \mathbf{W}\eta$ biztosan a $H \doteq \operatorname{lin}(\mathbf{W})$ altérre koncentrálódó normális eloszlású vektor. A ζ kovariancia mátrixa éppen

$$\begin{aligned} \mathbf{M}\left(\zeta \cdot \zeta^T\right) &\doteq \mathbf{M}\left(\mathbf{W}\eta \cdot (\mathbf{W}\eta)^T\right) = \mathbf{M}\left(\mathbf{W}\eta \cdot \eta^T \mathbf{W}^T\right) = \\ &= \mathbf{W} \cdot \mathbf{M}\left(\eta \cdot \eta^T\right) \cdot \mathbf{W}^T = \mathbf{W}\mathbf{W}^T \doteq \mathbf{S}, \end{aligned}$$

tehát a $\boldsymbol{\xi}$ és a ζ karakterisztikus függvénye megegyezik így a $\boldsymbol{\xi}$ és a $\zeta \doteq \mathbf{W}\eta$ eloszlása megegyezik, tehát $\mathbf{P}(\boldsymbol{\xi} \in H) = 1$.

Többdimenziós normális eloszlás

.

Example

A kétdimenziós normális eloszlás sűrűségfüggvénye.

Tegyük fel, hogy a (ξ, η) pár kétdimenziós normális eloszlású. Tegyük fel, hogy az eloszlás nem elfajuló. Legyen

$$\mathbf{S} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

a kovariancia mátrix.

$$\begin{aligned} |\mathbf{S}| &= \sigma_1^2\sigma_2^2 - r^2(\sigma_1\sigma_2)^2 = \left(1 - r^2\right)\sigma_1^2\sigma_2^2. \\ \mathbf{S}^{-1} &= \frac{1}{\left(1-r^2\right)\sigma_1^2\sigma_2^2}\begin{pmatrix} \sigma_2^2 & -r\sigma_1\sigma_2 \\ -r\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}. \end{aligned}$$

.

Ebből a sűrűségfüggvény

$$f(x,y) = \frac{1}{\sigma_1\sigma_2\sqrt{1-r^2}\left(\sqrt{2\pi}\right)^2}\times$$

$$\times\exp\left(-\frac{1}{2\left(1-r^2\right)}\left(\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2r\frac{\left(x-m_1\right)}{\sigma_1}\frac{\left(y-m_2\right)}{\sigma_2} + \left(\frac{y-m_2}{\sigma_1}\right)^2\right)\right)$$

Többdimenziós normális eloszlás

.

Example

A korrelálatlanságból csak akkor következik a függetlenség, ha az együttes eloszlás normális.

Legyen $\xi = N(0,1)$ és $\eta = \xi^2$. $\mathbf{E}(\eta) = 1$, így

$$\mathbf{E}((\xi - \mathbf{E}(\xi))(\eta - \mathbf{E}(\eta))) = \mathbf{E}(\xi(\eta - \mathbf{1})) =$$
$$= \mathbf{E}(\xi^3) - \mathbf{E}(\xi) = 0$$

Többdimenziós normális eloszlás

.

Example

Két normális eloszlású valószínűségi változó összege nem feltétlenül normális!

Természetesen az előzőek alapján ez csak akkor fordulhat elő, ha az összeadandók együttes eloszlása nem normális. Legyenek F_1 illetve F_2 kétdimenziós normális eloszlások. Tekintsük az $F \overset{\circ}{=} \left(F_1 + F_2\right)/2$ eloszlásfüggvényt. Tegyük fel, hogy mind a két eloszlásra nézve mind a két peremeloszlás várható értéke nulla, szórása egy, és az F_1-hez tartozó r_1 korrelációs együttható különbözik az F_2-höz tartozó r_2 együtthatótól.

.

Az F-nek van sűrűségfüggvénye, amely a két sűrűségfüggvény $1/2$ súllyal vett kombinációja.

$$\begin{aligned} f(x,y) &= \frac{1}{4\pi\sqrt{1-r_1^2}}\exp\left(-\frac{1}{2\left(1-r_1^2\right)}\left(x^2-2r_1xy+y^2\right)\right)+ \\ &+\frac{1}{4\pi\sqrt{1-r_2^2}}\exp\left(-\frac{1}{2\left(1-r_2^2\right)}\left(x^2-2r_2xy+y^2\right)\right). \end{aligned}$$

Többdimenziós normális eloszlás

.

A definíciókból világos, hogy az x és y szerinti peremeloszlásokra

$$
\begin{aligned}
g(x) &\overset{\circ}{=} \int_{\mathbb{R}} f(x,y)\,dy = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \\
h(y) &\overset{\circ}{=} \int_{\mathbb{R}} f(x,y)\,dx = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right).
\end{aligned}
$$

Ha a (ξ, η) eloszlása F, akkor a ξ és az η standard normális eloszlású. Számoljuk ki a $\xi + \eta$ összeg eloszlását!

$$
\int_{\mathbb{R}} f(z-y,y)\,dy = \frac{1}{2}\frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma_1^2}\right) + \frac{1}{2}\frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma_2^2}\right)
$$

ahol $\sigma_1 \overset{\circ}{=} \sqrt{2(1+r_1)}$, $\sigma_2 \overset{\circ}{=} \sqrt{2(1+r_2)}$. Mivel $r_1 \neq r_2$, ezért $\sigma_1 \neq \sigma_2$, ezért az összeg eloszlása nem normális.