Valószínűségszámítás

4. előadás

Arató Miklós

2017.10.02.

Tartalomjegyzék

Definíciók

Definíció: $\xi : \Omega \to \mathbb{R}$ **valószínűségi változó**, ha minden $x \in \mathbb{R}$ számra $\{w : \xi(w) < x\} \in \mathcal{A}$.

Definíciók

Definíció: $\xi : \Omega \to \mathbb{R}$ **valószínűségi változó**, ha minden $x \in \mathbb{R}$ számra $\{w : \xi(w) < x\} \in \mathcal{A}$.
Definíció: A ξ valószínűségi változó **diszkrét**, ha értékkészlete véges vagy megszámlálható, azaz léteznek olyan x_i valós számok és A_i teljes eseményrendszer, hogy $\xi = \sum_k x_k \cdot \chi_{A_k}$.
Definíció: A ξ valószínűségi változó **eloszlásfüggvénye** $F_\xi(x) = P(\xi < x)$, ahol $x \in \mathbb{R}$.
Definíció: A ξ valószínűségi változó abszolút folytonos eloszlású, ha $F_\xi(x) = P(\xi < x) = \int_{-\infty}^{x} f(s)\, ds$ minden $x \in \mathbb{R}$, ahol f véges sok ponttól eltekintve folytonos
f: sűrűségfüggvény

Példák

1. indikátor
2. Binomiális, $B(n, p)$
3. Geometriai (Pascal)
4. Hipergeometrikus
5. Poisson
6. Negatív binomiális
7. Egyenletes eloszlás, $E(a, b)$
8. λ-exponenciális
9. Gamma

Standard normális eloszlás

A ξ valószínűségi változó standard normális eloszlású, ha sűrűségfüggvénye:

$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \ (x \in \mathbb{R})$.

$$\left(\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \ dt\right)^2 = \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \ dt \cdot \int_{-\infty}^{\infty} e^{-\frac{s^2}{2}} \ ds =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{t^2+s^2}{2}} \ dt \ ds \stackrel{(\varphi,r)}{=} \int_0^{2\pi} d\varphi \cdot \int_0^{\infty} e^{-\frac{r^2}{2}} \cdot r \ dr =$$

$$2\pi \cdot \left[-e^{-\frac{r^2}{2}}\right]_0^{\infty} = 2\pi \Rightarrow . \ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} \ dt = 1$$

eloszlásfüggvény: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \ dt$.

Jelölése: $N(0,1)$.

Normális eloszlás

ξ valószínűségi változó standard normális eloszlású $\Rightarrow \eta = m + \sigma\xi$ normális eloszlású.

Eloszlásfüggvénye:

$P(m + \sigma\xi < x) = P(\xi < \frac{x-m}{\sigma}) = \Phi(\frac{x-m}{\sigma})$.

A sűrűségfüggvény: $f_{m+\sigma\xi}(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}$.

Ez a normális eloszlás m és σ^2 paraméterekkel, jelölése: $N(m, \sigma^2)$.

Fordítva, ha $\eta \sim N(m, \sigma^2)$, akkor $\frac{\eta - m}{\sigma} \sim N(0, 1)$.

Haranggörbe

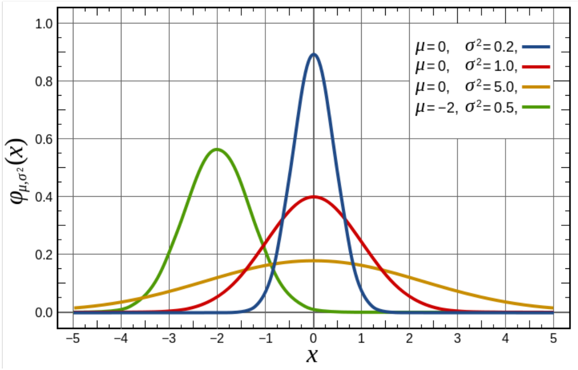

Figure: Normális sűrűségfüggvények

Táblázat

A standard normális eloszlásfüggvény táblázata

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,5000	0,45	0,6736	0,90	0,8159	1,35	0,9115	1,80	0,9641	2,50	0,9938
0,01	0,5040	0,46	0,6772	0,91	0,8186	1,36	0,9131	1,81	0,9649	2,52	0,9941
0,02	0,5080	0,47	0,6808	0,92	0,8212	1,37	0,9147	1,82	0,9656	2,54	0,9945
0,03	0,5120	0,48	0,6844	0,93	0,8238	1,38	0,9162	1,83	0,9664	2,56	0,9948
0,04	0,5160	0,49	0,6879	0,94	0,8264	1,39	0,9177	1,84	0,9671	2,58	0,9951
0,05	0,5199	0,50	0,6915	0,95	0,8289	1,40	0,9192	1,85	0,9678	2,60	0,9953
0,06	0,5239	0,51	0,6950	0,96	0,8315	1,41	0,9207	1,86	0,9686	2,62	0,9956
0,07	0,5279	0,52	0,6985	0,97	0,8340	1,42	0,9222	1,87	0,9693	2,64	0,9959
0,08	0,5319	0,53	0,7019	0,98	0,8365	1,43	0,9236	1,88	0,9699	2,66	0,9961
0,09	0,5359	0,54	0,7054	0,99	0,8389	1,44	0,9251	1,89	0,9706	2,68	0,9963
0,10	0,5398	0,55	0,7088	1,00	0,8413	1,45	0,9265	1,90	0,9713	2,70	0,9965
0,11	0,5438	0,56	0,7123	1,01	0,8438	1,46	0,9279	1,91	0,9719	2,72	0,9967
0,12	0,5478	0,57	0,7157	1,02	0,8461	1,47	0,9292	1,92	0,9726	2,74	0,9969
0,13	0,5517	0,58	0,7190	1,03	0,8485	1,48	0,9306	1,93	0,9732	2,76	0,9971
0,14	0,5557	0,59	0,7224	1,04	0,8508	1,49	0,9319	1,94	0,9738	2,78	0,9973
0,15	0,5596	0,60	0,7257	1,05	0,8531	1,50	0,9332	1,95	0,9744	2,80	0,9974
0,16	0,5636	0,61	0,7291	1,06	0,8554	1,51	0,9345	1,96	0,9750	2,82	0,9976
0,17	0,5675	0,62	0,7324	1,07	0,8577	1,52	0,9357	1,97	0,9756	2,84	0,9977
0,18	0,5714	0,63	0,7357	1,08	0,8599	1,53	0,9370	1,98	0,9761	2,86	0,9979
0,19	0,5753	0,64	0,7389	1,09	0,8621	1,54	0,9382	1,99	0,9767	2,88	0,9980
0,20	0,5793	0,65	0,7422	1,10	0,8643	1,55	0,9394	2,00	0,9772	2,90	0,9981
0,21	0,5832	0,66	0,7454	1,11	0,8665	1,56	0,9406	2,02	0,9783	2,92	0,9983
0,22	0,5871	0,67	0,7486	1,12	0,8686	1,57	0,9418	2,04	0,9793	2,94	0,9984

Meghatározás

Definíció: A $\xi_1, \ldots, \xi_n$ valószínűségi változók **függetlenek**, ha $P(\xi_1 < x_1, \ldots, \xi_n < x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i)$ minden $x_1, \ldots, x_n$-re.

Definíció: A $\xi_1, \xi_2, \ldots$ valószínűségi változók **függetlenek**, ha minden n-re $\xi_1, \ldots, \xi_n$ függetlenek.

Legyenek $\xi_1, \ldots, \xi_n$ függetlenek és g_i Borel-mérhető. Ekkor $g_1(\xi_1), \ldots, g_n(\xi_n)$ is függetlenek.

Legyenek $\xi_1, \ldots, \xi_n$ diszkrétek. Ekkor pontosan akkor függetlenek, ha $P(\xi_1 = x_1, \ldots, \xi_n = x_n) = \prod_{i=1}^{n} P(\xi_i = x_i)$ minden x_i-re.

Diszkrét konvolúciós formula

Legyenek ξ és η függetlenek, értékkészletük pedig $\{x_k\}$ és $\{y_l\}$.

$$P(\xi+\eta=z)=P\Big(\bigcup_{x_k+y_l=z}\{\xi=x_k,\eta=y_l\}\Big)=$$

$$\sum_{x_k+y_l=z} P(\xi=x_k,\eta=y_l)=\sum_{x_k+y_l=z} P(\xi=x_k)\cdot P(\eta=y_l).$$

Binomiálisak konvolúciójaa

Legyenek $\xi \sim B(n_1, p)$ és $\eta \sim B(n_2, p)$ függetlenek. Ekkor $\xi + \eta \sim B(n_1 + n_2, p)$.

Ugyanis $\xi + \eta$ értékkészlete $0, 1, \ldots, n_1 + n_2$, így

$$P(\xi + \eta = k) = \sum_{l=0}^{k} P(\xi = l) \cdot P(\eta = k - l) =$$

$$\sum_{l=\max\{k-n_2,0\}}^{\min\{k,n_1\}} \binom{n_1}{l} \cdot p^l \cdot (1-p)^{n_1-l} \cdot \binom{n_2}{k-l} \cdot p^{k-l} \cdot (1-p)^{n_2-k+l} =$$

$$\sum_{l=\max\{k-n_2,0\}}^{\min\{k,n_1\}} \binom{n_1}{l} \cdot \binom{n_2}{k-l} \cdot p^k \cdot (1-p)^{n_1+n_2-k} =$$

$p^k \cdot (1-p)^{n_1+n_2-k} \cdot \binom{n_1+n_2}{k}$, azaz $\xi + \eta \sim B(n_1 + n_2, p)$.

Binomiálisak konvolúciójaa

Legyenek $\xi \sim B(n_1, p)$ és $\eta \sim B(n_2, p)$ függetlenek. Ekkor $\xi + \eta \sim B(n_1 + n_2, p)$.
Ugyanis $\xi + \eta$ értékkészlete $0, 1, \ldots, n_1 + n_2$, így

$$P(\xi + \eta = k) = \sum_{l=0}^{k} P(\xi = l) \cdot P(\eta = k - l) =$$

$$\sum_{l=\max\{k-n_2,0\}}^{\min\{k,n_1\}} \binom{n_1}{l} \cdot p^l \cdot (1-p)^{n_1-l} \cdot \binom{n_2}{k-l} \cdot p^{k-l} \cdot (1-p)^{n_2-k+l} =$$

$$\sum_{l=\max\{k-n_2,0\}}^{\min\{k,n_1\}} \binom{n_1}{l} \cdot \binom{n_2}{k-l} \cdot p^k \cdot (1-p)^{n_1+n_2-k} =$$

$p^k \cdot (1-p)^{n_1+n_2-k} \cdot \binom{n_1+n_2}{k}$, azaz $\xi + \eta \sim B(n_1 + n_2, p)$.
egyszerűbben is kiszámítható: legyenek $X_1, X_2, \ldots$ független, azonos eloszlású p-indikátorok, ekkor $X_1 + \ldots + X_n \sim B(n, p)$, $X_{n+1} + \ldots + X_{n+m} \sim B(m, p)$, és így $X_1 + \ldots + X_{n+m} \sim B(n + m, p)$.

Poisson eset

Legyenek $\xi \sim \lambda$-Poisson és $\eta \sim \mu$-Poisson függetlenek. Ekkor $\xi + \eta \sim (\lambda + \mu)$-Poisson.

Ugyanis $P(\xi + \eta = k) = \sum_{l=0}^{k} P(\xi = l) \cdot P(\eta = k - l) =$

$$\sum_{l=0}^{k} \frac{\lambda^l \cdot e^{-\lambda}}{l!} \cdot \frac{\mu^{k-l} \cdot e^{-\mu}}{(k-l)!} = \frac{e^{-(\lambda+\mu)}}{k!} \cdot \sum_{l=0}^{k} \binom{k}{l} \cdot \lambda^l \cdot \mu^{k-l} = \frac{e^{-(\lambda+\mu)}}{k!} \cdot (\lambda+\mu)^k,$$

$\Rightarrow$

$(\lambda + \mu)$ paraméterű Poisson-eloszlást kapunk.

Konvolúciós formula

Legyenek ξ és η független, abszolút folytonos valószínűségi változók. Ekkor $\xi + \eta$ is abszolút folytonos eloszlású, és sűrűségfüggvénye

$$f_{\xi+\eta}(x) = \int_{-\infty}^{+\infty} f_\xi(x-y) \cdot f_\eta(y)\; dy = \int_{-\infty}^{+\infty} f_\xi(y) \cdot f_\eta(x-y)\; dy.$$

Exponenciálisak konvolúciója

$\xi_1, \dots, \xi_n$ független λ-exponenciális valószínűségi változók
$\eta_n = \xi_1 + \dots + \xi_n \Rightarrow$.

Állítás: η_n sűrűségfüggvénye $g_n(x) = \begin{cases} 0 & x \le 0 \\ \frac{x^{n-1} \cdot \lambda^n \cdot e^{-\lambda x}}{(n-1)!} & x > 0 \end{cases}$.

Bizonyítás: n-re vonatkozó teljes indukció
$n = 1$ rendben . Tegyük fel, hogy n-ig igaz az állítás. $\Rightarrow (n+1)$-re:

$$g_{n+1}(x) = f_{(\xi_1+\dots+\xi_n+\xi_{n+1})}(x) = \int_{-\infty}^{\infty} \underbrace{f_{\xi_1+\dots+\xi_n}(x-y)}_{g_n(x-y)} \cdot \underbrace{f_{\xi_{n+1}}(y)}_{g_1(y)}\, dy =$$

$$= \int_0^{\infty} \frac{(x-y)^{n-1} \lambda^n e^{-\lambda(x-y)}}{(n-1)!} \lambda e^{-\lambda y}\, dy =$$

$$\frac{\lambda^{n+1} e^{-\lambda x}}{n!} \underbrace{\int_0^x n(x-y)^{n-1} dy} = \frac{x^n \lambda^{n+1}}{n!} e^{-\lambda x}$$

"Buszok száma"

Olyan autóbuszjáratot, ahol a buszok követési ideje egymástól független, azonos λ-exponenciális eloszlású.
ξ_1 : az első busz beérkezési ideje, ξ_2 : az első és a második busz érkezése közötti idő, stb. N busz érkezik a $[0, t)$-ben.

$$P(N = n) = P(N \geq n) - P(N \geq n+1) = P(\eta_n < t) - P(\eta_{n+1} < t) =$$

$$\int_0^t \frac{x^{n-1}\lambda^n e^{-\lambda x}}{(n-1)!} dx - \int_0^t \frac{x^n \lambda^{n+1} e^{-\lambda x}}{n!} dx =$$

$$\frac{\lambda^n}{n!} \left\{ \int_0^t n x^{n-1} e^{-\lambda x} dx - \underbrace{\int_0^t x^n \lambda e^{-\lambda x} dx}_{[x^n e^{-\lambda x}]_0^t - \int_0^t n x^{n-1} e^{-\lambda x} dx} \right\} = \frac{(\lambda t)^n e^{-\lambda t}}{n!},$$

Meghatározás és tulajdonságok

Definíció: $\xi \geq 0$ **diszkrét**, ekkor $E\xi = \sum_k x_k \cdot P(\xi = x_k)$.

Definíció: ξ **diszkrét**, ekkor $E\xi = E\xi^{-} E\xi^{-}$, ha $\min(E\xi^{+}, E\xi^{-}) < \infty$.

Tulajdonságok:

1) $E(c \cdot \xi) = c \cdot E\xi$.
2) Ha $\xi \leq \eta \Rightarrow E\xi \leq E\eta$.
3) Ha létezik $E\xi$, akkor $|E\xi| \leq E|\xi|$.
4) Ha létezik $E\xi$, $E\eta$, és $E\xi + E\eta$ értelmes, akkor $E(\xi + \eta) = E\eta + E\xi$ is létezik.
5) Ha $P(\xi = 0) = 1 \Rightarrow E\xi = 0$.

Diszkrét példák

1) c konstans valószínűségi változó $\Rightarrow Ec = c$.

2) A indikátorárának várható értéke
$E\chi_A = 0 \cdot P(\chi_A = 0) + 1 \cdot P(\chi_A = 1) = P(A)$.

3) $\xi \sim B(n, p)$ **binomiális** eloszlású valószínűségi változó. Ekkor

$$E\xi = \sum_{k=0}^{n} k\binom{n}{k} p^k (1-p)^{n-k} =$$

$$np \cdot \underbrace{\sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1}(1-p)^{n-(k-1)}}_{=(p+(1-p))^{n-1}=1} = np.$$

Diszkrét példák

1) c konstans valószínűségi változó $\Rightarrow Ec = c$.

2) A indikátorárának várható értéke
$E\chi_A = 0 \cdot P(\chi_A = 0) + 1 \cdot P(\chi_A = 1) = P(A)$.

3) $\xi \sim B(n,p)$ **binomiális** eloszlású valószínűségi változó. Ekkor

$$E\xi = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} =$$

$$np \cdot \underbrace{\sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1}(1-p)^{n-(k-1)}}_{=(p+(1-p))^{n-1}=1} = np.$$

$\xi = \eta_1 + \ldots + \eta_n$ (független indikátorok összege)$\Rightarrow$
$E\xi = E\eta_1 + \ldots + E\eta_n = np$.

Diszkrét példák folytatás

4) $\xi \sim \lambda$-**Poisson**, ekkor

$$E\xi = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} = \lambda \cdot \underbrace{\sum_{k=1}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{k-1}}{(k-1)!}}_{=1} = \lambda.$$

5) Visszatevés nélkül húzzunk a dobozból. Ekkor a ξ valószínűségi változó **hipergeometrikus** eloszlású, és várható értéke:

$E\xi = \sum_{k=0}^{\min(n,M)} k \cdot \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$. Az egyszerűbb kiszámítás céljából legyen $\xi = \eta_1 + \ldots + \eta_M$, ahol $\eta_i := 1$, ha az i-edik piros golyót kihúztuk, különben pedig 0. Ekkor $P(\eta_i = 1) = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}$, így

$E\xi = n \cdot \frac{M}{N}$.

Meghatározás és tulajdonságok

Definíció: $\xi \geq 0$ **abszolút folytonos eloszlású**, ekkor $E\xi = \int_{\mathbb{R}} x \cdot f_\xi(x)\ dx$.

Definíció: ξ **abszolút folytonos eloszlású**, ekkor $E\xi = E\xi^- E\xi^-$, ha $\min(E\xi^+, E\xi^-) < \infty$.

$Eg(\xi) = \int_{\mathbb{R}} g(x) \cdot f_\xi(x)\ dx$

$\xi \geq 0 \Rightarrow E\xi = \int_0^\infty (1 - F_\xi(y))\ dy$

ξ, η függetlenek, $E|\xi|$, $E|\eta|$ végesek. $\Rightarrow E|\xi \cdot \eta|$ is véges és $E(\xi \cdot \eta) = E(\xi) \cdot E(\eta)$.

Abszolút folytonos eloszlású példák

1) $\xi \sim E(a,b)$ esetén $E\xi = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{a+b}{2}$.

2) λ-**exponenciális** eloszlás várható értéke $E\xi = \int_0^\infty e^{-\lambda y}\, dy = \frac{1}{\lambda}$.

3) $\xi \sim N(0,1)$ esetén $E\xi = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}\, dx = 0$, hiszen a sűrűségfüggvény szimmetrikus, így az integrálban egy páratlan függvény szerepel (továbbá az integrál konvergens, mert elég nagy x-re $x \cdot e^{-\frac{x^2}{2}}$ felülről becsülhető az e^{-x} függvénnyel). Általánosan pedig $m + \sigma\xi \sim N(m, \sigma^2)$ esetén $E(m + \sigma\xi) = m + \sigma \cdot E\xi = m$.

Meghatározás és Teljes várható érték tétel

Definíció: $P(A) > 0$, ξ diszkrét, ekkor ξ **feltételes várható értéke** az A feltételre nézve:
$E(\xi|A) = \sum_k x_k \cdot P(\xi = x_k|A)$.

Teljes várható érték tétel: ξ diszkrét valószínűségi változó véges várható értékkel és $A_1, A_2, \ldots$ teljes eseményrendszer, ahol $0 < P(A_i)$. Ekkor $E\xi = \sum_n E(\xi|A_n) \cdot P(A_n)$.

Wald-azonosság

$X_1, X_2, \dots$ független, azonos eloszlású valószínűségi változók, amelyekre létezik EX_i, és N egy tőlük független pozitív, egészértékű valószínűségi változó.

Ekkor $E(X_1 + \dots X_N) = EX_1 \cdot EN$.

$P(X_1 + \dots + X_N = y | N = n) = \frac{P(X_1 + \dots + X_n = y, N = n)}{P(N = n)} =$

$P(X_1 + \dots + X_n = y) \Rightarrow E(X_1 + \dots + X_N = y | N = n) = n \cdot EX_1 \Rightarrow$

$E(X_1 + \dots + X_N) = \sum_{n=1}^{\infty} E(X_1 + \dots + X_N | N = n) \cdot P(N = n) =$

$\sum_{n=1}^{\infty} n \cdot EX_1 \cdot P(N = n) = EX_1 \cdot EN$

Geometriai eloszlás

$P(\eta = k) = (1-p)^{k-1} \cdot p$, ahol $k \geq 1$.

$E\eta = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \cdot p$.

A : az első kísérlet sikeres

$E\eta = E(\eta | A) \cdot P(A) + E(\eta | \overline{A}) \cdot P(\overline{A}) \Rightarrow$

$E\eta = 1 \cdot p + (1 + E\eta) \cdot (1-p) \Rightarrow E\eta = \frac{1}{p}$

Várhatóan mikor lesz meg mind a 24 törpünk?

$\frac{24}{24} + \frac{24}{23} + \ldots + \frac{24}{2} + \frac{24}{1} = 90,623$

szimmetrikus bolyongás

ξ : lépések száma n-ből 0-ba $(n \geq 0)$, $v(n) := E\xi$.
A : az első lépésben jobbra lépünk $\Rightarrow E\xi = E(\xi|A) \cdot \frac{1}{2} + E(\xi|\overline{A}) \cdot \frac{1}{2}$
$E(\xi|A) = 1 + v(n+1)$ és $E(\xi|\overline{A}) = 1 + v(n-1) \Rightarrow$
$2v(n) = 2 + v(n+1) + v(n-1)$
$v(0) = 0$
$v(n+1) - v(n) = v(n) - v(n-1) - 2 = \ldots = v(1) - v(0) - 2n.$
$\Rightarrow$
$v(n) = v(n) - v(n-1) + v(n-1) + v(n-2) + \ldots + v(1) - v(0) + v(0)$
$\Rightarrow n \cdot v(1) - n(n-1) = n(v(1) - (n-1)) \geq 0 \Rightarrow v(1) \geq n-1$
minden n-re $\Rightarrow v(1) = +\infty \Rightarrow$
szimmetrikus bolyongásnál várhatóan végtelen sok lépésben térünk vissza a kiindulási pontba.