Statistika

Vanja Wagner

3. Osnove vjerojatnosti

Neprekidna distribucija.

Promotrimo sada slučajne varijable koje mogu poprimiti beskonačno **neprebrojivo** mnogo vrijednosti. U takvu klasu spada slučajna varijabla koja mjeri (nezaokruženu) visinu slučajno odabranog studenta Biologije.

Neprekidna distribucija.

Promotrimo sada slučajne varijable koje mogu poprimiti beskonačno **neprebrojivo** mnogo vrijednosti. U takvu klasu spada slučajna varijabla koja mjeri (nezaokruženu) visinu slučajno odabranog studenta Biologije.

Za slučajnu varijablu X kažemo da je **neprekidna** ukoliko postoji funkcija f_X takva da je

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f_X(t)dt.$$

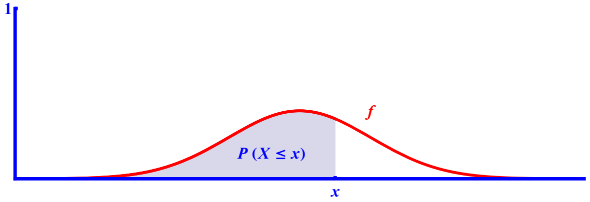
Neprekidna distribucija.

Promotrimo sada slučajne varijable koje mogu poprimiti beskonačno **neprebrojivo** mnogo vrijednosti. U takvu klasu spada slučajna varijabla koja mjeri (nezaokruženu) visinu slučajno odabranog studenta Biologije.

Za slučajnu varijablu X kažemo da je **neprekidna** ukoliko postoji funkcija f_X takva da je

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f_X(t)dt.$$

Gornja vjerojatnost je zadana površinom ispod krivulje funkcije f_X:



Vjerojatnosna funkcija gustoće

Funkcija f_X se naziva **vjerojatnosna funkcija gustoće** slučajne varijable X i zadovoljava sljedeća svojstva:

Vjerojatnosna funkcija gustoće

Funkcija f_X se naziva **vjerojatnosna funkcija gustoće** slučajne varijable X i zadovoljava sljedeća svojstva:

- $f(x) \geqslant 0$ - vjerojatnost je uvijek pozitivna

Vjerojatnosna funkcija gustoće

Funkcija f_X se naziva **vjerojatnosna funkcija gustoće** slučajne varijable X i zadovoljava sljedeća svojstva:

- $f(x) \geqslant 0$ - vjerojatnost je uvijek pozitivna
- $\displaystyle\int_{-\infty}^{\infty} f(t)dt = 1$ - vjerojatnost sigurnog događaja je 1.

Vjerojatnosna funkcija gustoće

Funkcija f_X se naziva **vjerojatnosna funkcija gustoće** slučajne varijable X i zadovoljava sljedeća svojstva:

- $f(x) \geqslant 0$ - vjerojatnost je uvijek pozitivna
- $\int_{-\infty}^{\infty} f(t) dt = 1$ - vjerojatnost sigurnog događaja je 1.

Očekivanje neprekidne slučajne varijable:

$$\mathbf{E}(X) = \int_{-\infty}^{\infty} t\, f(t)\, dt$$

Vjerojatnosna funkcija gustoće

Funkcija f_X se naziva **vjerojatnosna funkcija gustoće** slučajne varijable X i zadovoljava sljedeća svojstva:

- $f(x) \geqslant 0$ - vjerojatnost je uvijek pozitivna
- $\displaystyle\int_{-\infty}^{\infty} f(t)dt = 1$ - vjerojatnost sigurnog događaja je 1.

Očekivanje neprekidne slučajne varijable:

$$\mathbf{E}(X) = \int_{-\infty}^{\infty} t\, f(t)\, dt$$

Varijanca:

$$\mathrm{Var}(X) = \mathbf{E}(X - \mathbf{E}(X))^2 = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \int_{-\infty}^{\infty} t^2\, f(t)\, dt - (\mathbf{E}(X))^2.$$

Normalna distribucija - motivacija

Naš glavni primjer neprekidne distribucije bit će **normalna** ili **Gaussova** distribucija, odnosno distribucija kojoj graf funkcije gustoće odgovara tzv. Gaussovoj krivulji (krivulji zvona).

Normalna distribucija - motivacija

Naš glavni primjer neprekidne distribucije bit će **normalna** ili **Gaussova** distribucija, odnosno distribucija kojoj graf funkcije gustoće odgovara tzv. Gaussovoj krivulji (krivulji zvona).

Promotrimo kao motivaciju grafički prikaz distribucija **standardizirane** binomne slučajne varijable.

Normalna distribucija - motivacija

Naš glavni primjer neprekidne distribucije bit će **normalna** ili **Gaussova** distribucija, odnosno distribucija kojoj graf funkcije gustoće odgovara tzv. Gaussovoj krivulji (krivulji zvona).

Promotrimo kao motivaciju grafički prikaz distribucija **standardizirane** binomne slučajne varijable. Neka je $X \sim B(n, \frac{1}{2})$ i

$$Y = \frac{X - \mathbf{E}(X)}{\sigma(X)} = \frac{X - \frac{n}{2}}{\sqrt{\frac{n}{4}}}.$$

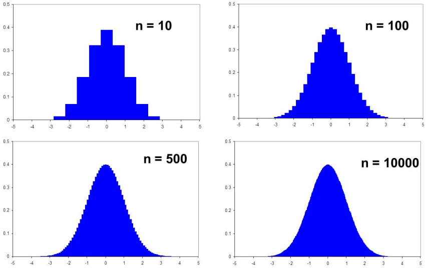
Normalna distribucija - motivacija

Naš glavni primjer neprekidne distribucije bit će **normalna** ili **Gaussova** distribucija, odnosno distribucija kojoj graf funkcije gustoće odgovara tzv. Gaussovoj krivulji (krivulji zvona).

Promotrimo kao motivaciju grafički prikaz distribucija **standardizirane** binomne slučajne varijable. Neka je $X \sim B(n, \frac{1}{2})$ i

$$Y = \frac{X - \mathbf{E}(X)}{\sigma(X)} = \frac{X - \frac{n}{2}}{\sqrt{\frac{n}{4}}}.$$

Promotrimo grafički prikaz distribucije od Y za $n = 10,\ 100,\ 500,\ 10000$.



2

Normalna distribucija - definicija

Distribucija neprekidne slučajne varijable je određena pripadnom vjerojatnosnom funkcijom gustoće.

Normalna distribucija - definicija

Distribucija neprekidne slučajne varijable je određena pripadnom vjerojatnosnom funkcijom gustoće. Tako kažemo da slučajna varijabla X ima **normalnu distribuciju** s parametrima $\mu \in \mathbf{R}$ i $\sigma^2 > 0$ ako je pripadna vjerojatnosna funkcija gustoće oblika

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbf{R}$$

i koristimo oznaku $X \sim N(\mu, \sigma^2)$.

Normalna distribucija - definicija

Distribucija neprekidne slučajne varijable je određena pripadnom vjerojatnosnom funkcijom gustoće. Tako kažemo da slučajna varijabla X ima **normalnu distribuciju** s parametrima $\mu \in \mathbf{R}$ i $\sigma^2 > 0$ ako je pripadna vjerojatnosna funkcija gustoće oblika

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbf{R}$$

i koristimo oznaku $X \sim N(\mu, \sigma^2)$. Parametri μ i σ^2 imaju direktnu interpretaciju:

$$\mathbf{E}(X) = \mu, \qquad \mathrm{Var}(X) = \sigma^2.$$

Jedinična normalna distribucija

Jedinična ili **standardna** normalna distribucija je normalna distribucija s parametrima

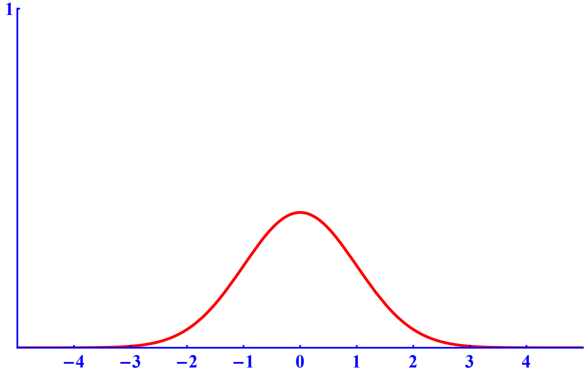
$$\mu = 0, \quad \sigma = 1.$$

Jedinična normalna distribucija

Jedinična ili **standardna** normalna distribucija je normalna distribucija s parametrima

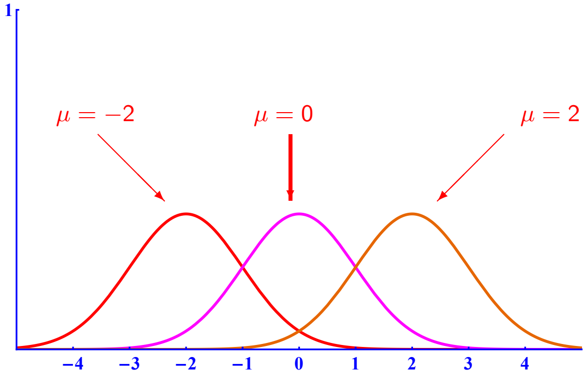
$$\mu = 0, \quad \sigma = 1.$$

Graf pripadne funkcije gustoće izgleda ovako:



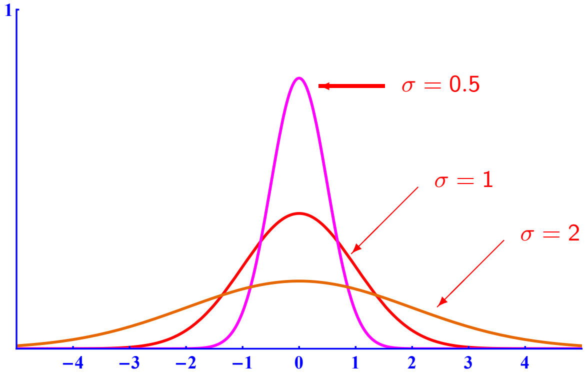
Kako se mijenja graf funkcije gustoće ako promijenimo parametar očekivanja μ?

Kako se mijenja graf funkcije gustoće ako promijenimo parametar očekivanja μ?
Promotrimo grafove funkcija gustoća za normalne distribucije s parametrima $\mu = -2, 0, 2$ i $\sigma = 1$:



Kako se mijenja graf funkcije gustoće ako promijenimo parametar varijance σ^2?

Kako se mijenja graf funkcije gustoće ako promijenimo parametar varijance σ^2? Promotrimo grafove funkcija gustoća za normalne distribucije s parametrima $\mu = 0$ i $\sigma = 0.5, 1, 2$.



Aproksimacija binomne distribucije normalnom distribucijom

Podsjetimo se našeg motivacijskog primjera - za $X \sim B(n, p)$ definiramo pripadnu standardiziranu slučajnu varijablu

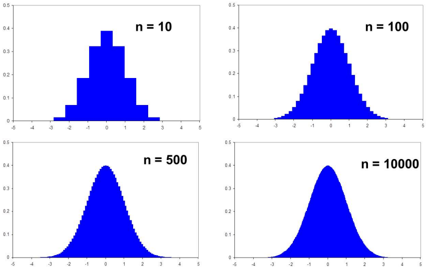
$$Y = \frac{X - \mathbf{E}(X)}{\sigma(X)} = \frac{X - np}{\sqrt{npq}}$$

Aproksimacija binomne distribucije normalnom distribucijom

Podsjetimo se našeg motivacijskog primjera - za $X \sim B(n,p)$ definiramo pripadnu standardiziranu slučajnu varijablu

$$Y = \frac{X - \mathbf{E}(X)}{\sigma(X)} = \frac{X - np}{\sqrt{npq}}$$

i promotrimo grafički prikaz distribucije od Y za $n = 10,\ 100,\ 500,\ 10000$:



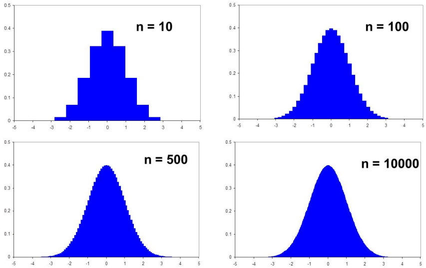
2

Aproksimacija binomne distribucije normalnom distribucijom

Podsjetimo se našeg motivacijskog primjera - za $X \sim B(n, p)$ definiramo pripadnu standardiziranu slučajnu varijablu

$$Y = \frac{X - \mathbf{E}(X)}{\sigma(X)} = \frac{X - np}{\sqrt{npq}}$$

i promotrimo grafički prikaz distribucije od Y za $n = 10,\ 100,\ 500,\ 10000$:



2

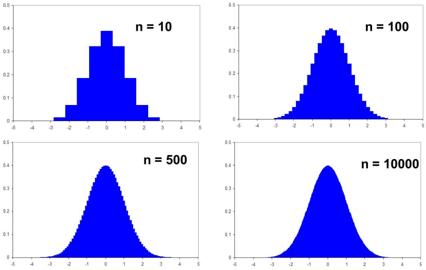
Uočimo da se distribucija ove standardizirane varijable Y stabilizira kada je n velik.

Aproksimacija binomne distribucije normalnom distribucijom

Podsjetimo se našeg motivacijskog primjera - za $X \sim B(n, p)$ definiramo pripadnu standardiziranu slučajnu varijablu

$$Y = \frac{X - \mathbf{E}(X)}{\sigma(X)} = \frac{X - np}{\sqrt{npq}}$$

i promotrimo grafički prikaz distribucije od Y za $n = 10,\ 100,\ 500,\ 10000$:



[2]

Uočimo da se distribucija ove standardizirane varijable Y stabilizira kada je n velik.

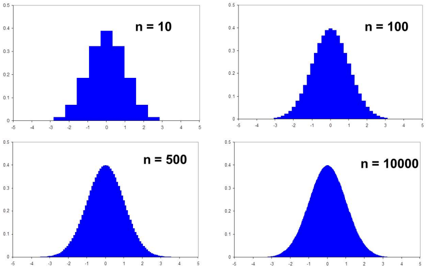
Za veliki n je distribucija standardizirane varijable Y slična distribuciji standardne normalne slučajne varijable.

Aproksimacija binomne distribucije normalnom distribucijom

Podsjetimo se našeg motivacijskog primjera - za $X \sim B(n,p)$ definiramo pripadnu standardiziranu slučajnu varijablu

$$Y = \frac{X - \mathbf{E}(X)}{\sigma(X)} = \frac{X - np}{\sqrt{npq}}$$

i promotrimo grafički prikaz distribucije od Y za $n = 10,\ 100,\ 500,\ 10000$:



[2]

Uočimo da se distribucija ove standardizirane varijable Y stabilizira kada je n velik.

Za veliki n je distribucija standardizirane varijable Y slična distribuciji standardne normalne slučajne varijable. $\longrightarrow$ izbor parametara $\mu = 0$ i $\sigma = 1$ ima smisla jer je $\mathbf{E}(Y) = 0,\ \mathrm{Var}(Y) = 1.$

Aproksimacija binomne distribucije normalnom distribucijom

Ovu opservaciju možemo matematički formulirati na sljedeći način: Ako je $X \sim B(n, p)$ za $n \in \mathbb{N}$ **velik** onda je pripadna standardizirana slučajna varijabla približno normalno distribuirana s parametrima $\mu = 0$ i $\sigma^2 = 1$,

Aproksimacija binomne distribucije normalnom distribucijom

Ovu opservaciju možemo matematički formulirati na sljedeći način: Ako je $X \sim B(n, p)$ za $n \in \mathbf{N}$ **velik** onda je pripadna standardizirana slučajna varijabla približno normalno distribuirana s parametrima $\mu = 0$ i $\sigma^2 = 1$, odnosno

$$\frac{X - np}{\sqrt{npq}} \approx N(0, 1).$$

Ovaj je rezultat poznat pod nazivom *de Moivre-Laplaceov* **centralni granični teorem**.

Aproksimacija binomne distribucije normalnom distribucijom

Ovu opservaciju možemo matematički formulirati na sljedeći način: Ako je $X \sim B(n, p)$ za $n \in \mathbb{N}$ **velik** onda je pripadna standardizirana slučajna varijabla približno normalno distribuirana s parametrima $\mu = 0$ i $\sigma^2 = 1$, odnosno

$$\frac{X - np}{\sqrt{npq}} \approx N(0, 1).$$

Ovaj je rezultat poznat pod nazivom *de Moivre-Laplaceov* **centralni granični teorem**.

Alternativno, možemo ga formulirati i kao:

$$X \approx N(np, npq)$$

i u ovom ga obliku najčešće i koristimo.

Funkcija distribucije standardne normalne razdiobe

Promotrimo sada kako računamo vjerojatnosti vezane uz normalne slučajne varijable. Za početak promotrimo $X \sim N(0,1)$. Kako bismo odredili vjerojatnost da je X najviše 2?

Funkcija distribucije standardne normalne razdiobe

Promotrimo sada kako računamo vjerojatnosti vezane uz normalne slučajne varijable. Za početak promotrimo $X \sim N(0,1)$. Kako bismo odredili vjerojatnost da je X najviše 2?

$$\mathbf{P}(X \leqslant 2) = F_X(2) = \int_{-\infty}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

Funkcija distribucije standardne normalne razdiobe

Promotrimo sada kako računamo vjerojatnosti vezane uz normalne slučajne varijable. Za početak promotrimo $X \sim N(0,1)$. Kako bismo odredili vjerojatnost da je X najviše 2?

$$\mathbf{P}(X \leqslant 2) = F_X(2) = \int_{-\infty}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

Ovaj se integral općenito ne može dobiti u zatvorenoj formi, već se računa **aproksimacijski**.

Funkcija distribucije standardne normalne razdiobe

Promotrimo sada kako računamo vjerojatnosti vezane uz normalne slučajne varijable. Za početak promotrimo $X \sim N(0,1)$. Kako bismo odredili vjerojatnost da je X najviše 2?

$$\mathbf{P}(X \leqslant 2) = F_X(2) = \int_{-\infty}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt.$$

Ovaj se integral općenito ne može dobiti u zatvorenoj formi, već se računa **aproksimacijski**. $\longrightarrow$ vrijednosti za F_X koji često označavamo s Φ su dane tablično.

Funkcija distribucije standardne normalne razdiobe

Promotrimo sada kako računamo vjerojatnosti vezane uz normalne slučajne varijable. Za početak promotrimo $X \sim N(0,1)$. Kako bismo odredili vjerojatnost da je X najviše 2?

$$\mathbf{P}(X \leqslant 2) = F_X(2) = \int_{-\infty}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

Ovaj se integral općenito ne može dobiti u zatvorenoj formi, već se računa **aproksimacijski**. $\longrightarrow$ vrijednosti za F_X koji često označavamo s Φ su dane tablično.

$\longrightarrow$ vrijednosti za F_X mogu se odrediti korištenjem statističkih alata (Excel, R, ...).

Funkcija distribucije standardne normalne razdiobe

Promotrimo sada kako računamo vjerojatnosti vezane uz normalne slučajne varijable. Za početak promotrimo $X \sim N(0,1)$. Kako bismo odredili vjerojatnost da je X najviše 2?

$$\mathbf{P}(X \leqslant 2) = F_X(2) = \int_{-\infty}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

Ovaj se integral općenito ne može dobiti u zatvorenoj formi, već se računa **aproksimacijski**. $\longrightarrow$ vrijednosti za F_X koji često označavamo s Φ su dane tablično.

$\longrightarrow$ vrijednosti za F_X mogu se odrediti korištenjem statističkih alata (Excel, R, ...).

Ako želimo odrediti $F_X(2)$ u Excelu pozivamo naredbu

```
= NORM.DIST(2, 0, 1, TRUE).
```

VAŽNO! Treći parametar u ovoj naredbi je standardna devijacija σ, a ne varijanca σ^2!

Neka svojstva normalne razdiobe

- Ako je X normalno distribuirana slučajna varijabla, onda je i $aX + b$ također normalno distribuirana slučajna varijabla, za sve $a, b \in \mathbf{R}$, $a \neq 0$.

Neka svojstva normalne razdiobe

- Ako je X normalno distribuirana slučajna varijabla, onda je i $aX + b$ također normalno distribuirana slučajna varijabla, za sve $a, b \in \mathbf{R}$, $a \neq 0$.
- Neka je $X \sim N(0, 1)$ i $Y = aX + b$. Tada je

$$Y \sim N(b, a^2).$$

Neka svojstva normalne razdiobe

- Ako je X normalno distribuirana slučajna varijabla, onda je i $aX + b$ također normalno distribuirana slučajna varijabla, za sve $a, b \in \mathbf{R}$, $a \neq 0$.
- Neka je $X \sim N(0,1)$ i $Y = aX + b$. Tada je

$$Y \sim N(b, a^2).$$

- Korištenjem ovog rezultata dobivamo da za $X \sim N(\mu, \sigma^2)$ vrijedi

$$\frac{X - \mu}{\sigma} \sim N(0,1).$$

Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Primjer. Ako slučajna varijabla X ima normalnu razdiobu $N(3, 4)$, kolika je vjerojatnost da će vrijednost slučajne varijable biti manja od 5?

Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Primjer. Ako slučajna varijabla X ima normalnu razdiobu $N(3,4)$, kolika je vjerojatnost da će vrijednost slučajne varijable biti manja od 5?

Rješenje. $\mu = 3, \quad \sigma^2 = 4, \quad \sigma = 2$

Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Primjer. Ako slučajna varijabla X ima normalnu razdiobu $N(3,4)$, kolika je vjerojatnost da će vrijednost slučajne varijable biti manja od 5?

Rješenje. $\mu = 3, \quad \sigma^2 = 4, \quad \sigma = 2$

$$P(X \leqslant 5) = ?$$

Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Primjer. Ako slučajna varijabla X ima normalnu razdiobu $N(3, 4)$, kolika je vjerojatnost da će vrijednost slučajne varijable biti manja od 5?

Rješenje. $\mu = 3, \quad \sigma^2 = 4, \quad \sigma = 2$

$$P(X \leqslant 5) = ?$$

`= NORM.DIST(5, 3, 2, TRUE).`

Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Primjer. Ako slučajna varijabla X ima normalnu razdiobu $N(3,4)$, kolika je vjerojatnost da će vrijednost slučajne varijable biti manja od 5?

Rješenje. $\mu = 3, \quad \sigma^2 = 4, \quad \sigma = 2$

$$P(X \leqslant 5) = ?$$

```
= NORM.DIST(5, 3, 2, TRUE).
```

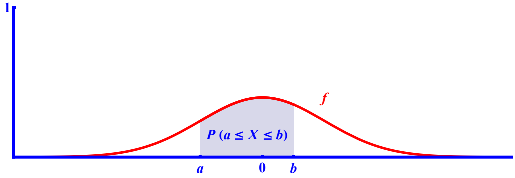
```
0.8413447
```

Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Kako izračunati $P(a \leqslant X \leqslant b)$?

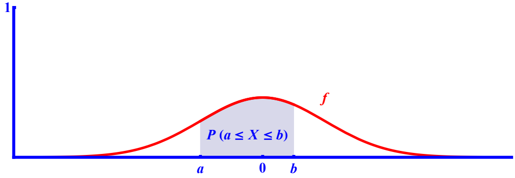
Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Kako izračunati $P(a \leqslant X \leqslant b)$?



Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

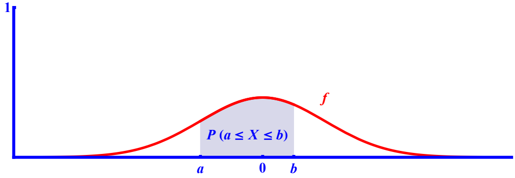
Kako izračunati $P(a \leqslant X \leqslant b)$?



$$P(a \leqslant X \leqslant b) = P(X \leqslant b) - P(X \leqslant a) = F_X(b) - F_X(a).$$

Računanje vjerojatnosti vezanih uz normalnu distribuciju - nastavak

Kako izračunati $P(a \leqslant X \leqslant b)$?



$$P(a \leqslant X \leqslant b) = P(X \leqslant b) - P(X \leqslant a) = F_X(b) - F_X(a).$$

Napomena. Uočite da je kod neprekidnih distribucija $P(X = a) = 0$ za svaki a.

Primjer. Kolika je vjerojatnost da će slučajna varijabla s normalnom razdiobom $N(0, 1)$ poprimiti vrijednost između -1 i 2?

Primjer. Kolika je vjerojatnost da će slučajna varijabla s normalnom razdiobom $N(0,1)$ poprimiti vrijednost između -1 i 2?

Rješenje. $\mu = 0, \quad \sigma^2 = 1, \quad \sigma = 1$

Primjer. Kolika je vjerojatnost da će slučajna varijabla s normalnom razdiobom $N(0,1)$ poprimiti vrijednost između -1 i 2?

Rješenje. $\mu = 0, \quad \sigma^2 = 1, \quad \sigma = 1$

$$P(-1 \leqslant X \leqslant 2) = P(X \leqslant 2) - P(X \leqslant -1) = F_X(2) - F_X(-1).$$

```
> = NORM.DIST(2, 0, 1, TRUE)
0.9772499
```

Primjer. Kolika je vjerojatnost da će slučajna varijabla s normalnom razdiobom $N(0,1)$ poprimiti vrijednost između -1 i 2?

Rješenje. $\mu = 0, \quad \sigma^2 = 1, \quad \sigma = 1$

$$P(-1 \leqslant X \leqslant 2) = P(X \leqslant 2) - P(X \leqslant -1) = F_X(2) - F_X(-1).$$

```
> = NORM.DIST(2, 0, 1, TRUE)
0.9772499
> = NORM.DIST(-1, 0, 1, TRUE)
0.1586553
```

Primjer. Kolika je vjerojatnost da će slučajna varijabla s normalnom razdiobom $N(0,1)$ poprimiti vrijednost između -1 i 2?

Rješenje. $\mu = 0, \quad \sigma^2 = 1, \quad \sigma = 1$

$$P(-1 \leqslant X \leqslant 2) = P(X \leqslant 2) - P(X \leqslant -1) = F_X(2) - F_X(-1).$$

```
> = NORM.DIST(2, 0, 1, TRUE)
0.9772499
> = NORM.DIST(-1, 0, 1, TRUE)
0.1586553
```

$$\begin{aligned} P(-1 \leqslant X \leqslant 2) &= P(X \leqslant 2) - P(X \leqslant -1) = \\ &= 0.977250 - 0.158655 = 0.818595 \end{aligned}$$

Aproksimacija binomne distribucije normalnom - nastavak

Kako koristimo ranije spomenuti rezultat gdje za velike n ($np \geqslant 5$ i $nq \geqslant 5$) distribuciju binomne slučajne varijable $X \sim B(n, p)$ možemo aproksimirati normalnom distribucijom,

$$X \approx N(np, npq)?$$

Aproksimacija binomne distribucije normalnom - nastavak

Kako koristimo ranije spomenuti rezultat gdje za velike n ($np \geqslant 5$ i $nq \geqslant 5$) distribuciju binomne slučajne varijable $X \sim B(n, p)$ možemo aproksimirati normalnom distribucijom,

$$X \approx N(np, npq)?$$

Neka je $F = F_{N(np,npq)}$ funkcija distribucije normalne razdiobe $N(np, npq)$. Tada

$$\mathbf{P}(X \leqslant b) \approx F(b) \text{ ili } F(b + \tfrac{1}{2}),$$
$$\mathbf{P}(a \leqslant X \leqslant b) \approx F(b) - F(a) \text{ ili } F(b + \tfrac{1}{2}) - F(a + \tfrac{1}{2}),$$

gdje je crveni izraz *bez korekcije* a plavi izraz *s korekcijom* po neprekidnosti. Obe aproksimacije su korisne, korekcija daje nešto bolju aproksimaciju uzimajući u obzir činjenicu da je binomna razdioba diskretna a normalna razdioba neprekidna.

Aproksimacija binomne distribucije normalnom - nastavak

Kako koristimo ranije spomenuti rezultat gdje za velike n ($np \geqslant 5$ i $nq \geqslant 5$) distribuciju binomne slučajne varijable $X \sim B(n, p)$ možemo aproksimirati normalnom distribucijom,

$$X \approx N(np, npq)?$$

Neka je $F = F_{N(np,npq)}$ funkcija distribucije normalne razdiobe $N(np, npq)$. Tada

$$\begin{aligned} \mathbf{P}(X \leqslant b) &\approx F(b) \text{ ili } F(b + \tfrac{1}{2}), \\ \mathbf{P}(a \leqslant X \leqslant b) &\approx F(b) - F(a) \text{ ili } F(b + \tfrac{1}{2}) - F(a + \tfrac{1}{2}), \end{aligned}$$

gdje je crveni izraz *bez korekcije* a plavi izraz *s korekcijom* po neprekidnosti.

Aproksimacija binomne distribucije normalnom - nastavak

Kako koristimo ranije spomenuti rezultat gdje za velike n ($np \geqslant 5$ i $nq \geqslant 5$) distribuciju binomne slučajne varijable $X \sim B(n, p)$ možemo aproksimirati normalnom distribucijom,

$$X \approx N(np, npq)?$$

Neka je $F = F_{N(np,npq)}$ funkcija distribucije normalne razdiobe $N(np, npq)$. Tada

$$\mathbf{P}(X \leqslant b) \approx F(b) \text{ ili } F(b + \tfrac{1}{2}),$$
$$\mathbf{P}(a \leqslant X \leqslant b) \approx F(b) - F(a) \text{ ili } F(b + \tfrac{1}{2}) - F(a + \tfrac{1}{2}),$$

gdje je crveni izraz *bez korekcije* a plavi izraz *s korekcijom* po neprekidnosti. Obe aproksimacije su korisne, korekcija daje nešto bolju aproksimaciju uzimajući u obzir činjenicu da je binomna razdioba diskretna a normalna razdioba neprekidna.