MTTTP1, luento 27.9.2018

7 TILASTOLLISEN PÄÄTTELYN PERUSTEITA

- Miten voidaan arvioida virheellisten komponenttien osuutta tuotannossa?
- Miten voidaan arvioida valmistajan kynttilöiden keskimääräistä palamisaikaa?
- Ovatko kaupungissa eri alueilla myynnissä olevien asuntojen keskineliöhinnat samoja?
- Riippuuko myytävän asunnon kunto sijainnista?
- Miten päättely populaatiosta otoksen perusteella tehdään?

Otos	Populaatio
otoskeskiarvo $\bar{x}$	populaation keskiarvo, odotusarvo μ
otosvarianssi s^2	populaation varianssi σ^2
otoshajonta s	populaation hajonta σ
%-osuus otoksessa p	%-osuus populaatiossa π

Tilastollisessa päättelyssä voidaan arvioida esim.

- odotusarvoa
- prosenttiosuutta
- kahden populaation odotusarvojen yhtäsuuruutta
- muuttujien riippumattomuutta

Otoksesta määritellyt $\bar{x}$, s^2, s, p ovat otossuureita, joiden käyttäytymistä voidaan arvioida todennäköisyysjakaumien avulla. Näitä jakaumia käytetään hyväksi päättelyssä.

7.1 Satunnaisilmiö ja tapahtuma

Esim. 7.1.1. Rahanheitto, nopanheitto, lottoaminen.

Satunnaisilmiö (satunnaiskoe)

useita tulosmahdollisuuksia, epävarmuus tuloksesta

Perusjoukko (E)

kaikki mahdolliset tulokset

Tapahtuma (A)

perusjoukon osajoukko

Esim. 7.1.2.

Rahanheitto

$E = \{\text{kruunu, klaava}\}$

tapahtumia

$A = \{\text{kruunu}\}$

$B = \{\text{klaava}\}$

Nopanheitto

$E = \{1, 2, 3, 4, 5, 6\}$

tapahtuma

$A = \{\text{saadaan parillinen}\} = \{2,4,6\}$

7.2 Klassinen todennäköisyys

Tapahtuman A todennäköisyys

$$P(A) = k/n$$

n satunnaisilmiön perusjoukon tulosten lukumäärä

k tapahtumaan A liittyvien tulosten lukumäärä

Esim. 7.2.1.

Rahanheitto

$A = \{kruunu\}$

$P(A) = 1/2$

Nopanheitto

$A = \{\text{saadaan parillinen}\} = \{2,4,6\}$

$P(A) = 3/6$

$B = \{1\}$, $P(B) = 1/6$

$D = \{\text{suurempi kuin } 4\} = \{5,6\}$, $P(D) = 2/6$

Tapahtumien A ja B riippumattomuus

7.3 Satunnaismuuttuja ja todennäköisyysjakauma

Esim. 7.3.1. Nopanheitto

X = saatu silmäluku

$P(X=1) = P(X=2) = \ldots = P(X=6) = 1/6$

Esim. 7.3.2. Heitetään kolikkoa neljä kertaa, X = klaavojen lukumäärä heittosarjassa

	lukumäärä		lukumäärä
Kl,Kl,Kl,Kl	4	Kr,Kl,Kl,Kr	2
Kr,Kl,Kl,Kl	3	Kl,Kr,Kl,Kr	2
Kl,Kr,Kl,Kl	3	Kr,Kl,Kr,Kl	2
Kl,Kl,Kr,Kl	3	Kl,Kr,Kr,Kr	1
Kl,Kl,Kl,Kr	3	Kr,Kl,Kr,Kr	1
Kl,Kl,Kr,Kr	2	Kr,Kr,Kl,Kr	1
Kr,Kr,Kl,Kl	2	Kr,Kr,Kr,Kl	1
Kl,Kr,Kr,Kl	2	Kr,Kr,Kr,Kr	0

$P(X=0) = 1/16$, $P(X=3) = 4/16$, $P(X=1) = 4/16$,
$P(X=4) = 1/16$, $P(X=2) = 6/16$

Esim. 7.3.4.

Kahden alkion otokset luvuista 1, 2, 3, 4, 5, 6 systemaattisella otannalla ovat {1, 4}, {2, 5}, {3, 6}, joista keskiarvot 2,5, 3,5 ja 4,5, joten

$$P(\bar{X}=2,5) = P(\bar{X}=3,5) = P(\bar{X}=4,5) = 1/3.$$

Satunnaismuuttuja

funktio, joka liittää yksikäsitteisen reaaliluvun jokaiseen tarkasteltavan satunnaisilmiön perusjoukon tulokseen

Diskreetin satunnaismuuttujan X todennäköisyysjakauma

$P(X=x_1) = p_1$, $P(X=x_2) = p_2$...

$p_1 + p_2 + ... = 1$

Jatkuvan satunnaismuuttujan X todennäköisyysjakauma

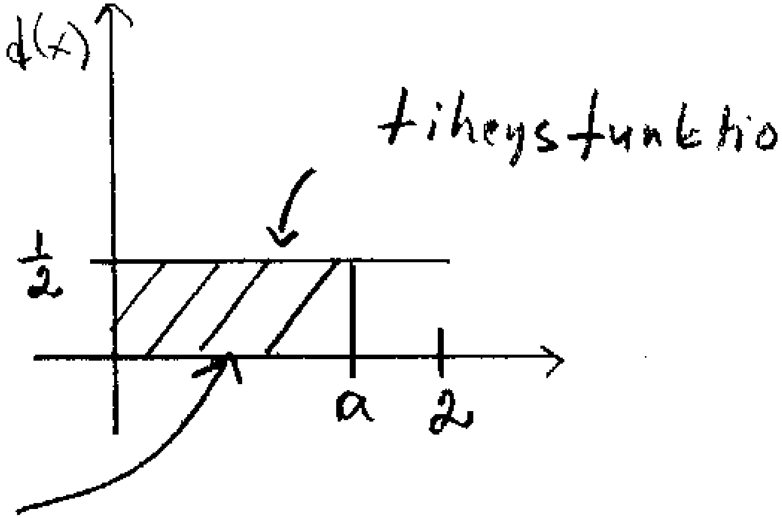
jatkuva funktio $f(x)$, jolle $f(x) \geq 0$ sekä $f(x)$:n ja x-akselin väliin jäävä pinta-ala on yksi.

Funktiota $f(x)$ kutsutaan tiheysfunktioksi.

Satunnaismuuttujan X kertymäfunktio

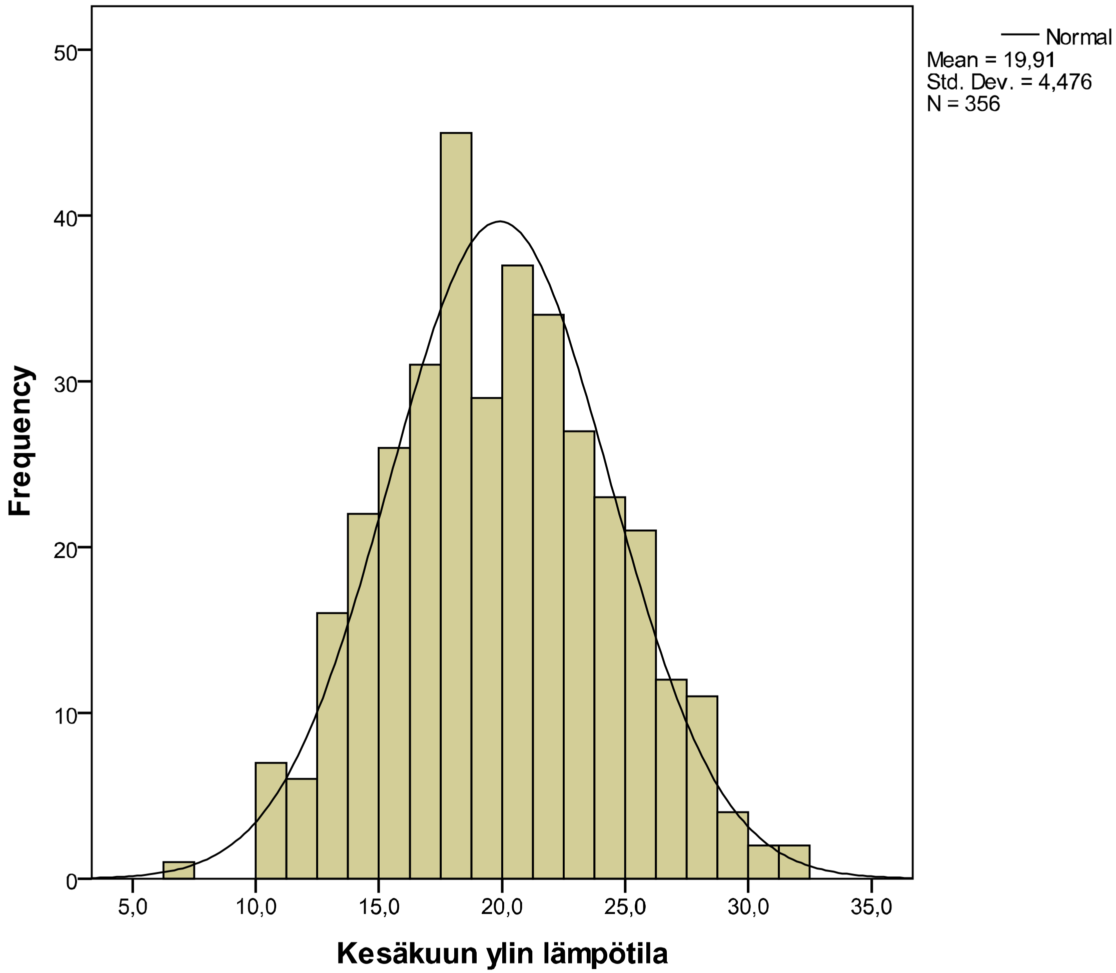
$F(x) = P(X \leq x)$.

Esim. 7.3.5. Esimerkki erään jatkuvan satunnaismuuttujan tiheys- ja kertymäfunktiosta



$$\text{Pinta-ala} = P(X \leq a) = \underbrace{F(a)}_{\text{kertymäfunktio}} = a \cdot \frac{1}{2}$$

Esim. Erään tiheysfunktion kuvaaja.



Todennäköisyysjakaumien tunnuslukuja

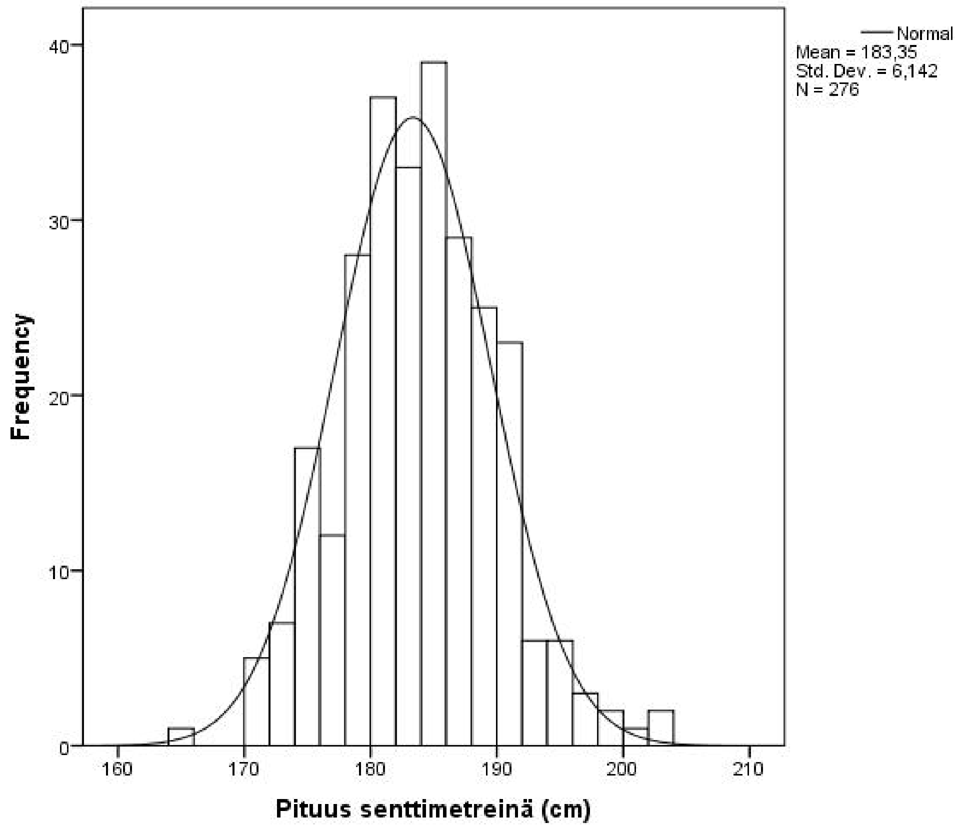
- odotusarvo $E(X) = \mu$
- varianssi $Var(X) = \sigma^2$, keskihajonta σ

Satunnaismuuttujien summat, erotukset, suhteet, jne. ovat myös satunnaismuuttujia.

Satunnaismuuttujien riippumattomuus määritellään vastaavalla tavalla kuin tapahtumien riippumattomuus.

7.4 Normaalijakauma

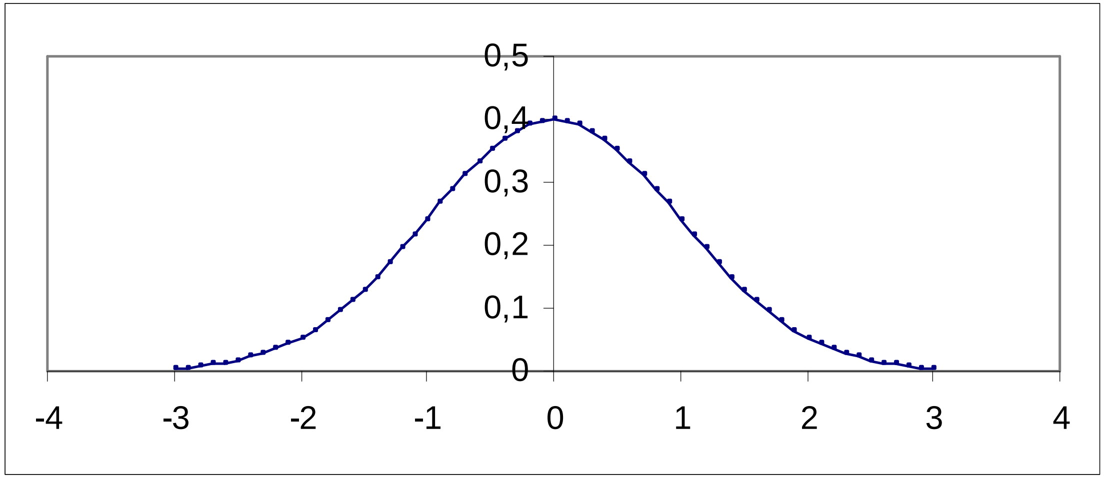
Esim. 7.4.1. Vaahteraliigan pelaajien pituusjakauma. Kuvaan on piirretty normaalijakauman, jonka odotusarvo 183,35 ja varianssi $6{,}142^2$, tiheysfunktio.



Normaalijakauma määritellään parametrein μ ja σ^2, merkitään $X \sim N(\mu, \sigma^2)$, tiheysfunktion kuvaajia, ks. https://fi.wikipedia.org/wiki/Normaalijakauma

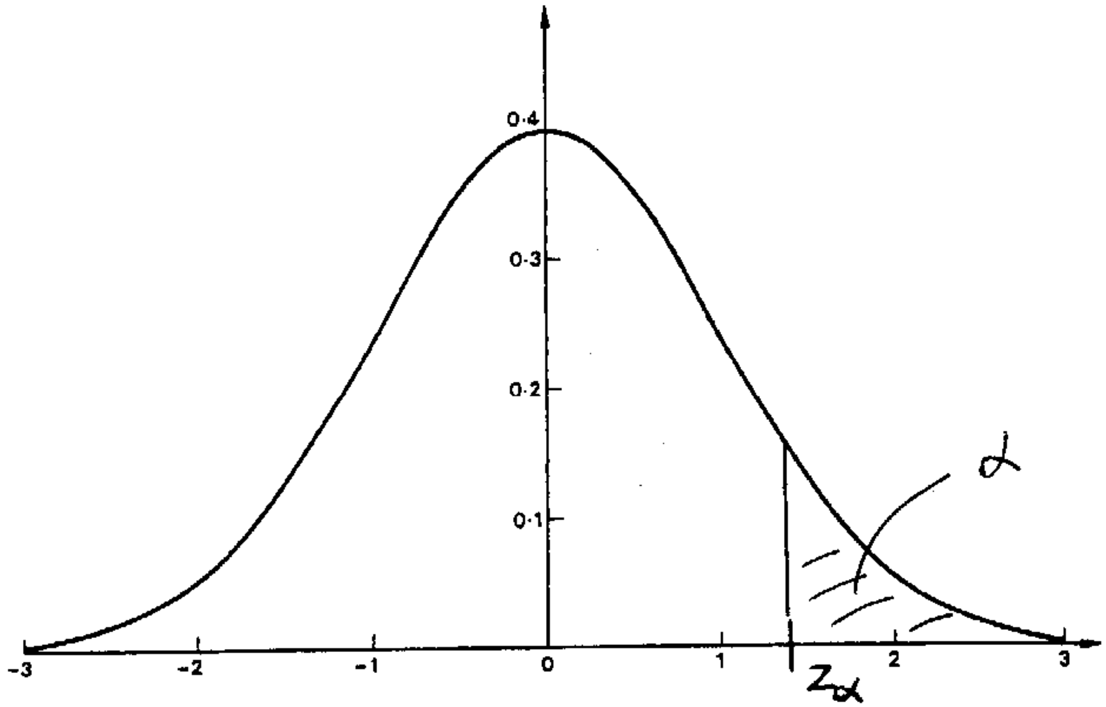
Jos odotusarvo on nolla ja varianssi yksi, kyseessä standardoitu normaalijakauma, merkitään $Z \sim N(0, 1)$. Tällöin $P(Z \leq z) = \Phi(z)$, standardoidun normaalijakauman kertymäfunktiota merkitään $\Phi(z)$:lla.

Esim. 7.4.2. N(0, 1) – jakauman tiheysfunktion kuvaaja

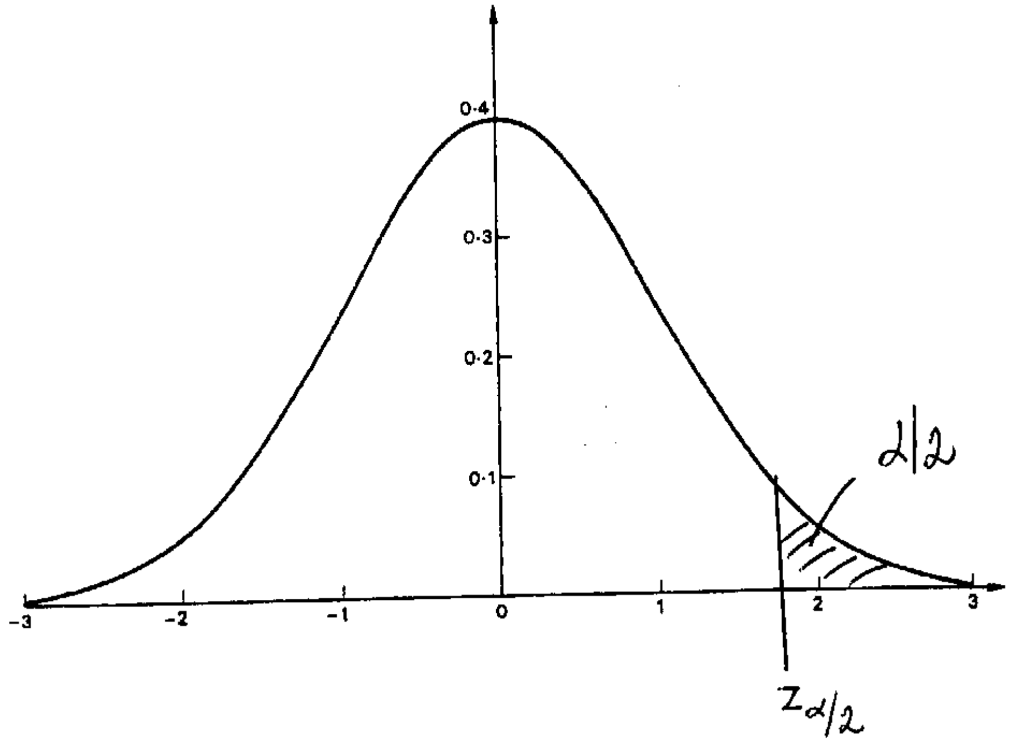


Standardoidun normaalijakauman kertymäfunktion Φ(z) arvoja taulukoitu, ks.
http://www.sis.uta.fi/tilasto/mtttp1/syksy2018/kaavat.pdf

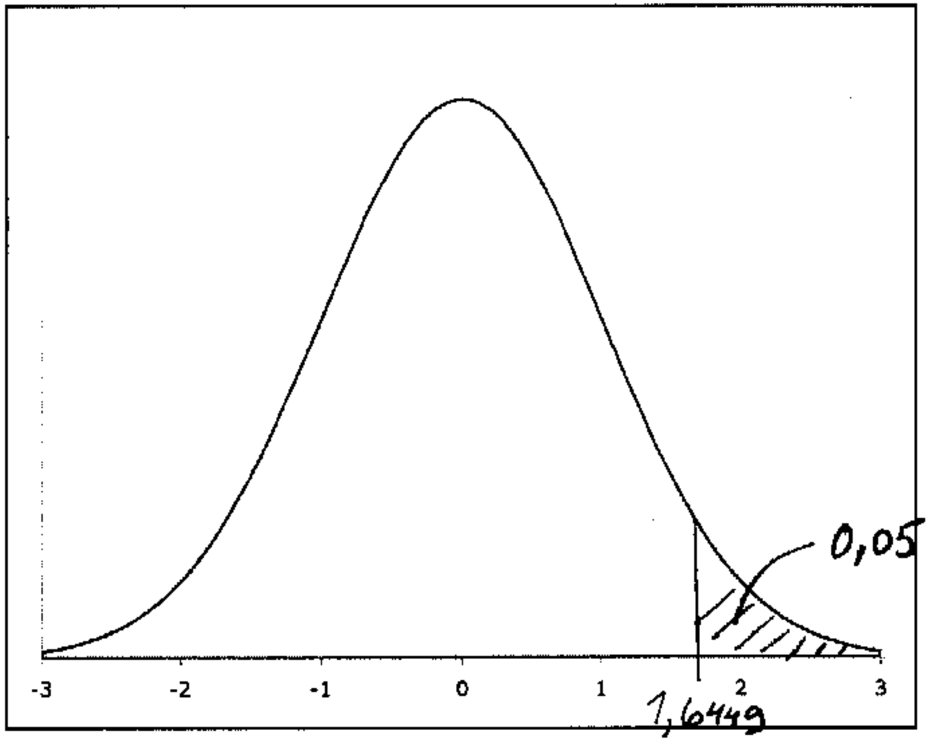
Olkoon Z ~ N(0, 1). Määritellään z_α siten, että $P(Z \geq z_\alpha) = \alpha$, graafisesti:



Samoin $z_{\alpha/2}$ siten, että $P(Z \geq z_{\alpha/2}) = \alpha/2$, graafisesti:

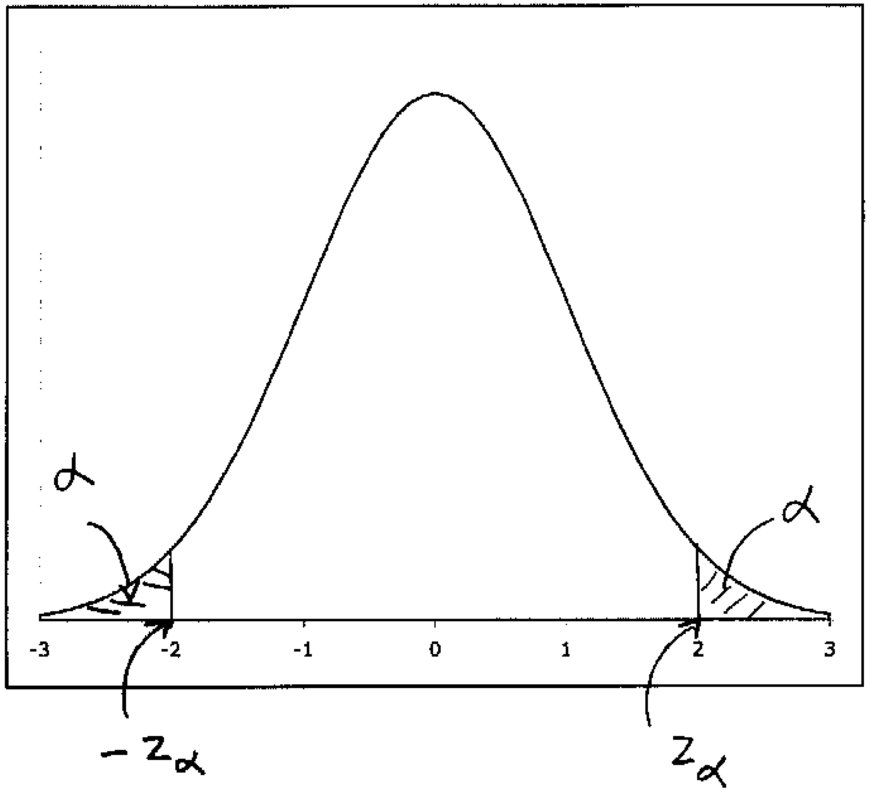


Esim. 7.4.3.



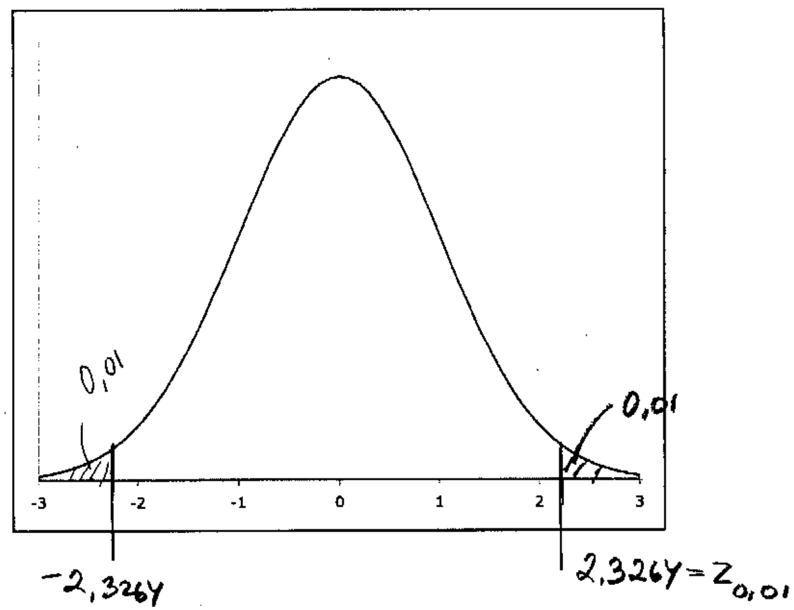
$z_{0,05} = 1,6449$
$z_{0,01} = 2,3264$
$z_{0,05/2} = z_{0,025} = 1,96$

Standardoitu normaalijakauman symmetrinen nollan suhteen



Esim. 7.4.4. Olkoon $Z \sim N(0, 1)$.

$P(Z \geq 2{,}3264) = 0{,}01$, $P(Z \leq -2{,}3264) = 0{,}01$



$P(Z \geq 1{,}96) = 0{,}025$, $P(Z \leq -1{,}96) = 0{,}025$

$P(Z \geq 1{,}6449) = 0{,}05$, $P(Z \leq -1{,}6449) = 0{,}05$