Johdatus todennäköisyyslaskentaan

Normaalijakaumasta johdettuja jakaumia

Normaalijakaumasta johdettuja jakaumia

Johdanto

χ^2-jakauma

***F*-jakauma**

***t*-jakauma**

Normaalijakaumasta johdettuja jakaumia: Mitä opimme? – 1/2

- Tutustumme tässä luvussa seuraaviin **normaalijakaumasta** (ks. lukua **Jatkuvia jakaumia**) **johdettuihin jakaumiin**:
  - **χ^2-jakauma**
  - ***F*-jakauma**
  - ***t*-jakauma**
- Tarkastelun kohteena ovat seuraavat χ^2-, *F*- ja *t*-jakaumien ominaisuudet:
  (i) **Jakauman määrittely**
  (ii) **Odotusarvo**, **varianssi** ja **standardipoikkeama**
  (iii) **Tiheysfunktion kuvaaja**

Normaalijakaumasta johdettuja jakaumia: Mitä opimme? – 2/2

- Lisäksi tarkastelemme **todennäköisyyksien määräämistä** χ^2-, *F*- ja *t*-jakaumista.
- Koska χ^2-, *F*- ja *t-jakaumien tiheysfunktioiden integraalifunktioita ei tunneta*, χ^2-, *F*- ja *t*-jakaumiin liittyvien todennäköisyyksien määräämisessä on käytettävä jotakin *numeerista menetelmää*.
- Siksi useimmissa tilastotieteen ja todennäköisyyslaskennan oppikirjoissa on valmiit *taulukot*, joissa on taulukoituna χ^2-, *F*- ja *t*-jakaumien kertymäfunktioiden arvoja ja niihin liittyviä todennäköisyyksiä.
- χ^2-, *F*- ja *t*-jakaumien *tiheysfunktioiden lausekkeet* johdetaan luvussa **Satunnaismuuttujien muunnokset ja niiden jakaumat.**

Normaalijakaumasta johdettuja jakaumia: Esitiedot

- Esitiedot: ks. seuraavia lukuja:

  Satunnaismuuttujat ja todennäköisyysjakaumat

  Kertymäfunktio

  Jakaumien tunnusluvut

  Jatkuvia jakaumia

Normaalijakaumasta johdettuja jakaumia: Lisätiedot

- χ^2-, F- ja t-jakaumien *tiheysfunktioiden lausekkeiden johtaminen* vaatii satunnaismuuttujan **2. potenssin** sekä *riippumattomien* satunnaismuuttujien **summan** ja **osamäärän jakaumien** määräämistä; ks. lisätietoja luvusta

  Satunnaismuuttujien muunnokset ja niiden jakaumat

- Huomautus:

  Tarkoitamme *satunnaismuuttujien riippumattomuudella* sitä, että yhdenkään satunnaismuuttujan saamat arvot eivät riipu siitä, mitä arvoja muut satunnaismuuttujat saavat; käsite täsmennetään luvussa **Kaksiulotteiset todennäköisyysjakaumat**.

Normaalijakaumasta johdettuja jakaumia

>> Johdanto

χ^2-jakauma

***F*-jakauma**

***t*-jakauma**

Johdanto

Avainsanat

- **χ^2-jakauma**
- ***F*-jakauma**
- **Jakaumien määritteleminen**
- ***t*-jakauma**

Jakaumien määritteleminen normaalijakauman avulla

- Useat tilastotieteen keskeiset todennäköisyysjakaumat voidaan *määritellä* normaalijakauman avulla.
- Tällaisia ovat esimerkiksi χ^2-, F- ja t-jakaumat, joilla on keskeinen rooli *otosjakaumien teoriassa*, *estimoinnissa* ja *testauksessa* (ks. esim. lukuja **Otos ja otosjakaumat**, **Estimointi** ja **Tilastollisten hypoteesien testaus**).
- Tarkastelemme seuraavien jakaumien määrittelemistä ja ominaisuuksia:
  - **χ^2-jakauma**
  - ***F*-jakauma**
  - ***t*-jakauma**

Normaalijakaumasta johdettuja jakaumia

Johdanto

>> χ^2-jakauma

***F*-jakauma**

***t*-jakauma**

χ^2-jakauma

Avainsanat

χ^2-jakauma
Normaalijakauma
Odotusarvo
Standardipoikkeama
Standardoitu normaalijakauma
Tiheysfunktio
Todennäköisyyksien määrääminen
χ^2-jakaumasta
Vapausasteet
Varianssi

χ^2-jakauman määritelmä 1/2

- Olkoot X_i , i = 1, 2, … , n *riippumattomia*, *standardoitua normaalijakaumaa* N(0,1) (ks. lukua **Jatkuvia jakaumia**) *noudattavia satunnaismuuttujia*.
- Tällöin

$$X_i \sim \mathrm{N}(0,1)\,,\; i = 1, 2, \ldots, n$$

$$X_1, X_2, \ldots, X_n \perp$$

χ^2-jakauman määritelmä 2/2

- Olkoon

$$X = \sum_{i=1}^{n} X_i^2$$

N(0,1)-jakautuneiden, *riippumattomien* satunnais-muuttujien X_i , $i = 1, 2, \ldots , n$ *neliösumma*.

- Tällöin satunnaismuuttuja X noudattaa **χ^2-jakaumaa** (Khiin neliö -jakaumaa) ***n*:llä vapausasteella**.
- Merkintä:

$$X \sim \chi^2(n)$$

χ^2-jakauman vapausasteet

- χ^2-jakauman vapausasteiden lukumäärä n viittaa *yhteenlaskettavien lukumäärään* χ^2-jakauman määrittelevässä neliösummassa.
- Vapausasteiden lukumäärä n on χ^2-jakauman muodon määräävä *parametri*.

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim \chi^2(n)$.
- **Odotusarvo**:

$$\mathrm{E}(X) = n$$

- **Varianssi** ja **standardipoikkeama**:

$$\mathrm{Var}(X) = \mathrm{D}^2(X) = 2n$$

$$\mathrm{D}(X) = \sqrt{2n}$$

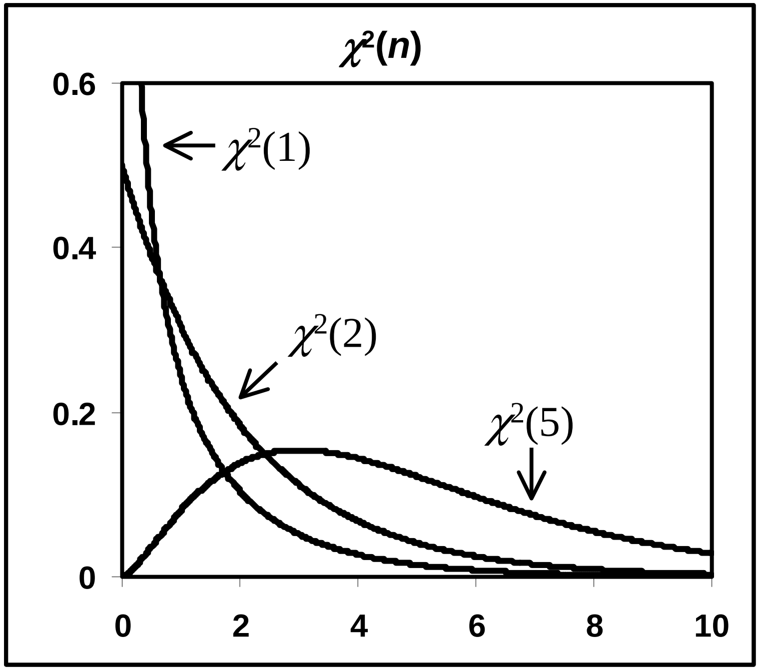
Tiheysfunktion kuvaaja

- Kuva oikealla esittää χ^2-jakauman
  $$\chi^2(n)$$
  tiheysfunktiota välillä [0, 10], kun vapausasteiden lukumäärällä n on seuraavat arvot:

  (i) $n = 1$

  (ii) $n = 2$

  (iii) $n = 5$
- Jakauman *odotusarvo*:
  $$\mathrm{E}(X) = n$$



χ^2-jakauma
Tiheysfunktion ja sen kuvaajan ominaisuuksia

- χ^2-jakauman tiheysfunktio $f(x)$ on *positiivinen* kaikille positiivisille argumentin arvoille:

  $$f(x) > 0 \ , \ x > 0$$

- Jos vapausasteiden lukumäärä

  $$n = 1, 2$$

  niin tiheysfunktio on *monotonisesti laskeva* kaikille $x \geq 0$.

- Jos vapausasteiden lukumäärä

  $$n \geq 3$$

  niin tiheysfunktio on *yksihuippuinen* ja sillä on *maksimi* jossakin pisteessä $x > 0$.

Todennäköisyyksien määrääminen χ^2-jakaumasta 1/2

- Todennäköisyydet voidaan määrätä χ^2-*jakaumasta* jakauman *kertymäfunktion* avulla.
- Olkoon $X \sim \chi^2(n)$.
- Olkoon satunnaismuuttujan X *kertymäfunktio*

$$F_{Chi}(x\,;n) = \Pr(X \le x)$$

- Huomautus 1:

  Merkinnällä $F_{Chi}(x\,;n)$ on haluttu korostaa χ^2-jakauman riippuvuutta sen vapausasteiden lukumäärästä n.

- Huomautus 2:

  χ^2-*jakauman tiheysfunktion integraalifunktiota ei tunneta*, joten χ^2-jakauman kertymäfunktion määräämiseen on käytettävä jotakin *numeerista menetelmää*.

Todennäköisyyksien määrääminen χ^2-jakaumasta 2/2

- *Kaikkien* χ^2-jakaumaan liittyvien tapahtumien todennäköisyydet saadaan todennäköisyyksistä

$$\Pr(X \le x) = F_{Chi}(x\,;\,n)$$

todennäköisyyslaskennan laskusääntöjen avulla.

- Esimerkiksi

$$\Pr(a \le X \le b) = F_{Chi}(b) - F_{Chi}(a)$$

Todennäköisyyksien määrääminen χ^2-jakaumasta: Taulukot 1/2

- χ^2-jakauman *taulukot* sisältävät tavallisesti *argumentin x arvoja* taulukoituna *useille vapausasteiden lukumäärille n*, mutta vain *muutamille kertymäfunktion* F_{Chi} *arvoille*.
- Siten taulukot mahdollistavat seuraavan tehtävän ratkaisemisen (taulukkokohtaisin rajoituksin):

  Määrää x, kun *todennäköisyys*

  $$\Pr(X \leq x) = F_{Chi}(x\,;\,n)$$

  on annettu.

Todennäköisyyksien määrääminen χ^2-jakaumasta: Taulukot 2/2

- Koska χ^2-jakaumaa käytetään tavallisesti *väliestimoinnin* tai *testauksen* yhteydessä, χ^2-jakauman taulukoihin on yleensä taulukoitu sellaisia argumentin x arvoja, jotka vastaavat todennäköisyyden

$$\Pr(X \le x) = F_{Chi}(x\,;\,n)$$

komplementtitodennäköisyyttä

$$p = \Pr(X \ge x) = 1 - F_{Chi}(x\,;\,n)$$

Todennäköisyyksien määrääminen χ^2-jakaumasta: Esimerkki

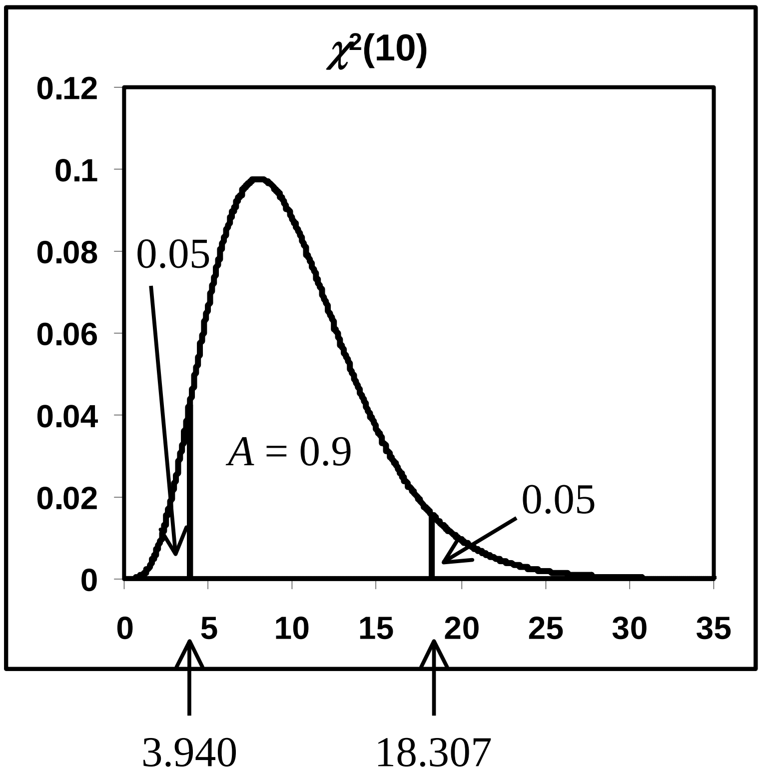
- Kuva oikealla esittää χ^2-jakauman

  $$\chi^2(10)$$

  tiheysfunktiota välillä [0, 35].
- χ^2-jakauman taulukoista saadaan:

  Alueen A pinta-ala

  $$= \Pr(3.940 \leq X \leq 18.307)$$
  $$= F_{Chi}(18.307;10)$$
  $$-F_{Chi}(3.940;10)$$
  $$= 0.95 - 0.05$$
  $$= 0.9$$



Todennäköisyyksien määrääminen χ^2-jakaumasta: Ohjelmat

- Olkoon $X \sim \chi^2(n)$.
- Monet *tietokoneohjelmat* mahdollistavat seuraavien tehtävien ratkaisemisen ilman χ^2-jakauman taulukoiden asettamia rajoituksia:

(i) Määrää todennäköisyys

$$\Pr(X \le x) = F_{Chi}(x\,;\,n)$$

kun x on annettu.

(ii) Määrää x, kun todennäköisyys

$$\Pr(X \le x) = F_{Chi}(x\,;\,n)$$

on annettu.

Normaalijakaumasta johdettuja jakaumia

Johdanto

χ^2-jakauma

>> *F*-jakauma

***t*-jakauma**

F-jakauma

Avainsanat

χ^2-jakauma
***F*-jakauma**
Normaalijakauma
Odotusarvo
Standardipoikkeama
Standardoitu normaalijakauma
Tiheysfunktio
Todennäköisyyksien määrääminen *F*-jakaumasta
Vapausasteet
Varianssi

F-jakauman määritelmä 1/2

- Olkoot Y_i , $i = 1, 2, \ldots , m$ ja X_i , $i = 1, 2, \ldots , n$ *riippumattomia, standardoitua normaalijakaumaa* N(0,1) (ks. lukua **Jatkuvia jakaumia**) *noudattavia satunnaismuuttujia*.
- Tällöin

$$Y_i \sim \mathrm{N}(0,1) \,,\ i = 1, 2, \ldots, m \,,\ X_i \sim \mathrm{N}(0,1) \,,\ i = 1, 2, \ldots, n$$

$$Y_1, Y_2, \ldots, Y_m, X_1, X_2, \ldots, X_n \perp$$

ja edelleen

$$Y = \sum_{i=1}^{m} Y_i^2 \sim \chi^2(m) \,,\ X = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$$

$$Y \perp X$$

F-jakauman määritelmä 2/2

- Olkoon

$$F = \frac{\frac{1}{m}Y}{\frac{1}{n}X} = \frac{n}{m} \cdot \frac{Y}{X}$$

jossa

$$Y \sim \chi^2(m)\,,\ X \sim \chi^2(n)\,,\ Y \perp X$$

- Tällöin satunnaismuuttuja F noudattaa **Fisherin *F*-jakaumaa *m*:llä ja *n*:llä vapausasteella**.
- Merkintä:

$$F \sim F(m, n)$$

F-jakauman vapausasteet

- *F*-jakauman vapausasteiden lukumääristä *ensimmäinen* (*m*) viittaa *yhteenlaskettavien lukumäärään F*-jakauman määrittelevän lausekkeen *osoittajassa*.
- *F*-jakauman vapausasteiden lukumääristä *toinen* (*n*) viittaa *yhteenlaskettavien lukumäärään F*-jakauman määrittelevän lausekkeen *nimittäjässä*.
- Vapausasteiden lukumäärät *m* ja *n* ovat *F*-jakauman muodon määrääviä *parametreja*.

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $F \sim F(m, n)$.
- **Odotusarvo**:

$$\mathrm{E}(F) = \frac{n}{n-2}\,,\ n > 2$$

- **Varianssi** ja **standardipoikkeama**:

$$\mathrm{Var}(F) = \mathrm{D}^2(F) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}\,,\ n > 4$$

$$\mathrm{D}(F) = \sqrt{\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}}\,,\ n > 4$$

F-jakauman ominaisuuksia

- Olkoon

$$F \sim F(m, n).$$

- Tällöin myös $1/F$ on F-jakautunut, mutta vapausastein n ja m:

$$\frac{1}{F} \sim F(n, m)$$

Tiheysfunktion kuvaaja

- Kuva oikealla esittää *F*-jakauman

  $F(m, n)$

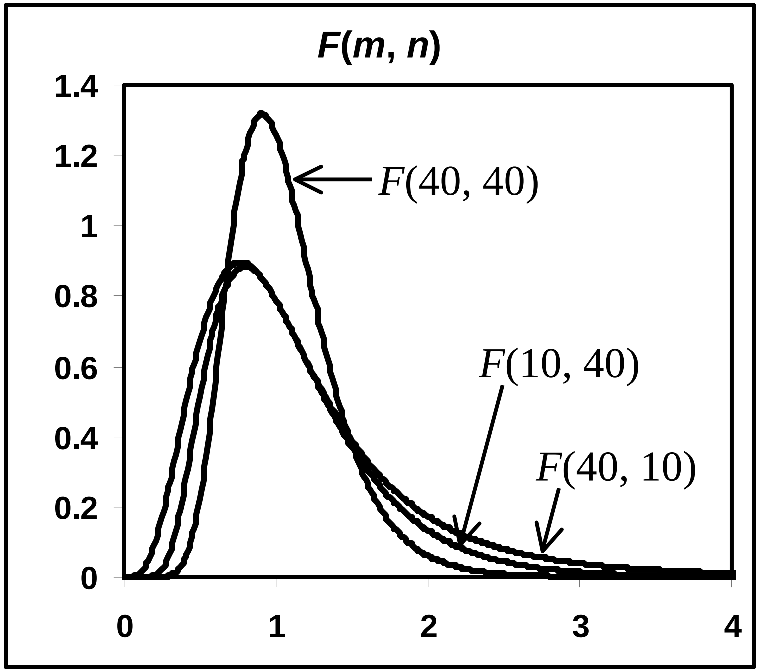
  tiheysfunktiota välillä [0, 5], kun vapausasteiden lukumäärillä m ja n on seuraavat arvot:

  (i) $m = 10, n = 40$

  (ii) $m = 40, n = 10$

  (iii) $m = 40, n = 40$
- Jakauman *odotusarvo*:

$$\mathrm{E}(F) = \frac{n}{n-2}, n > 2$$



F-jakauma

Tiheysfunktion ja sen kuvaajan ominaisuuksia

- *F*-jakauman tiheysfunktio $f(x)$ on *positiivinen* kaikille positiivisille argumentin arvoille:

$$f(x) > 0 \ , \ x > 0$$

- Jos *osoittajan* vapausasteiden lukumäärä

$$m = 1, 2$$

niin tiheysfunktio on *monotonisesti laskeva* kaikille $x \geq 0$.

- Jos *osoittajan* vapausasteiden lukumäärä

$$m \geq 3$$

niin tiheysfunktio on *yksihuippuinen* ja sillä on *maksimi* jossakin pisteessä $x > 0$.

Todennäköisyyksien määrääminen *F*-jakaumasta 1/2

- Todennäköisyydet voidaan määrätä *F-jakaumasta* jakauman *kertymäfunktion* avulla.
- Olkoon $F \sim F(m, n)$.
- Olkoon satunnaismuuttujan F *kertymäfunktio*

$$F_F(x\,;\,m,n) = \Pr(F \le x)$$

- Huomautus 1:

  Merkinnällä $F_F(x\,;\,m,n)$ on haluttu korostaa F-jakauman riippuvuutta sen vapausasteiden lukumääristä m ja n.

- Huomautus 2:

  F-jakauman tiheysfunktion integraalifunktiota ei tunneta, joten F-jakauman kertymäfunktion määräämiseen on käytettävä jotakin *numeerista menetelmää*.

Todennäköisyyksien määrääminen *F*-jakaumasta 2/2

- *Kaikkien F*-jakaumaan liittyvien tapahtumien todennäköisyydet saadaan todennäköisyyksistä

$$\Pr(F \leq x) = F_F(x\,;\, m, n)$$

todennäköisyyslaskennan laskusääntöjen avulla.

- Esimerkiksi

$$\Pr(a \leq F \leq b) = F_F(b) - F_F(a)$$

Todennäköisyyksien määrääminen *F*-jakaumasta: Taulukot 1/4

- *F*-jakauman *taulukot* sisältävät tavallisesti *argumentin x arvoja* taulukoituina *useille vapausasteiden lukumäärille m* ja *n*, mutta vain *muutamille kertymäfunktion* F_F *arvoille*.
- Siten taulukot mahdollistavat seuraavan tehtävän ratkaisemisen (taulukkokohtaisin rajoituksin):

  Määrää *x*, kun todennäköisyys

  $$\Pr(F \leq x) = F_F(x\,;\,m,\,n)$$

  on annettu.

Todennäköisyyksien määrääminen *F*-jakaumasta: Taulukot 2/4

- Koska F-jakaumaa käytetään tavallisesti *väliestimoinnin* tai *testauksen* yhteydessä, F-jakauman taulukoihin on yleensä taulukoitu sellaisia argumentin x arvoja, jotka vastaavat todennäköisyyden

  $$\Pr(F \leq x) = F_F(x\,;\,m,n)$$

  komplementtitodennäköisyyttä

  $$p = \Pr(F \geq x) = 1 - F_F(x\,;\,m,n).$$

Todennäköisyyksien määrääminen *F*-jakaumasta: Taulukot 3/4

- Monet *F*-jakauman taulukot sisältävät todennäköisyyksiä

$$p = \Pr(F \geq x) = 1 - F_F(x\,;\, m, n)$$

vastaavia argumentin arvoja vain, kun p on "*pieni*".

- "*Suuriin*" p:n arvoihin liittyvät argumentin x arvot saadaan tällöin käyttämällä hyväksi sitä, että $1/F \sim F(n, m)$.
- Olkoon

$$F_{m,n} \sim F(m, n) \ \text{ ja } \ p = \Pr(F_{m,n} \leq a)$$

$$F_{n,m} \sim F(n, m) \ \text{ ja } \ p = \Pr(F_{n,m} \geq b)$$

- Tällöin

$$a = \frac{1}{b}$$

Todennäköisyyksien määrääminen *F*-jakaumasta: Taulukot 4/4

- Oletukset:

$$F_{m,n} \sim F(m, n)$$
$$F_{n,m} \sim F(n, m)$$
$$p = \Pr(F_{m,n} \leq a)$$
$$= \Pr(F_{n,m} \geq b)$$

- Tällöin:

$$a = \frac{1}{b}$$

- Perustelu:

Todetaan ensin, että

$$p = \Pr(F_{m,n} \leq a)$$
$$= \Pr(1/F_{m,n} \geq 1/a)$$
$$= \Pr(F_{n,m} \geq 1/a)$$

Koska oletuksen mukaan

$$p = \Pr(F_{n,m} \geq b)$$

niin

$$b = 1/a$$

Todennäköisyyksien määrääminen *F*-jakaumasta: Esimerkki

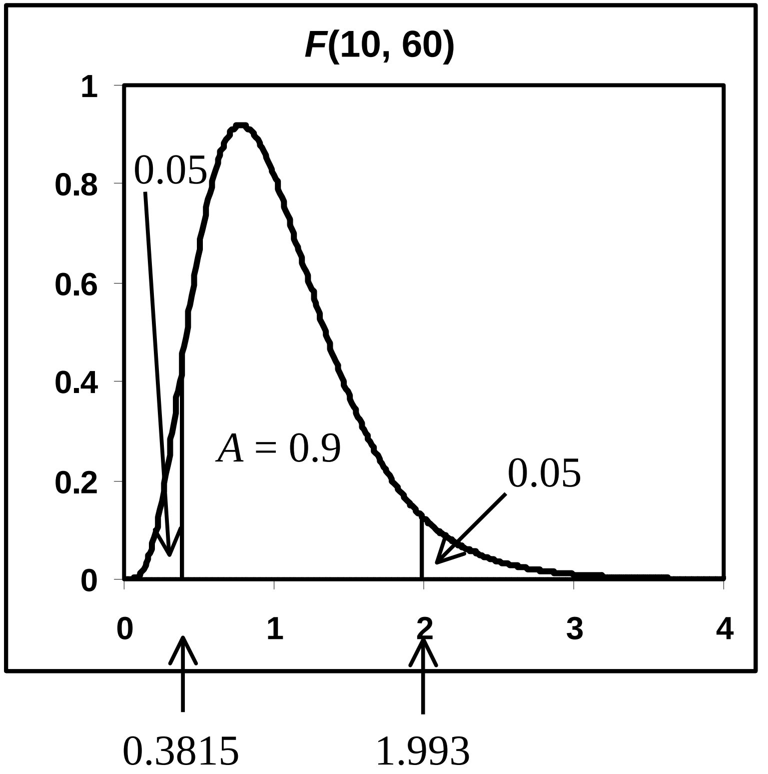
- Kuva oikealla esittää *F*-jakauman

  $F(10, 60)$

  tiheysfunktiota välillä [0, 4].
- *F*-jakauman taulukoista saadaan:

  Alueen A pinta-ala

  $$\begin{aligned}
  &= \Pr(0.3815 \le F \le 1.993) \\
  &= F_F(1.993; 10, 60) \\
  &\qquad - F_F(0.3815; 10, 60) \\
  &= 0.95 - 0.05 \\
  &= 0.9
  \end{aligned}$$



Todennäköisyyksien määrääminen *F*-jakaumasta: Ohjelmat

- Olkoon $F \sim F(m, n)$.
- Useat *tietokoneohjelmat* mahdollistavat seuraavien tehtävien ratkaisemisen ilman *F*-jakauman taulukoiden asettamia rajoituksia:

(i) Määrää todennäköisyys

$$\Pr(F \leq x) = F_F(x\ ;\ m, n)$$

kun x on annettu.

(ii) Määrää x, kun todennäköisyys

$$\Pr(F \leq x) = F_F(x\ ;\ m, n)$$

on annettu.

Normaalijakaumasta johdettuja jakaumia

Johdanto

χ^2-jakauma

***F*-jakauma**

>> *t*-jakauma

t-jakauma

Avainsanat

- **χ^2-jakauma**
- ***F*-jakauma**
- **Normaalijakauma**
- **Odotusarvo**
- **Standardipoikkeama**
- **Standardoitu normaalijakauma**
- ***t*-jakauma**
- **Tiheysfunktio**
- **Todennäköisyyksien määrääminen *t*-jakaumasta**
- **Vapausasteet**
- **Varianssi**

t-jakauman määritelmä 1/2

- Olkoot Y ja X_i , i = 1, 2, … , n *riippumattomia*, *standardoitua normaalijakaumaa* N(0,1) (ks. lukua **Jatkuvia jakaumia**) *noudattavia satunnaismuuttujia*.
- Tällöin

$$Y \sim \mathrm{N}(0,1)\,,\; X_i \sim \mathrm{N}(0,1)\,,\; i = 1, 2, \ldots, n$$

$$Y, X_1, X_2, \ldots, X_n \perp$$

ja edelleen

$$X = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$$

$$Y \perp X$$

t-jakauman määritelmä 2/2

- Olkoon

$$t = \frac{Y}{\sqrt{\frac{1}{n}X}}$$

jossa

$$Y \sim \mathrm{N}(0,1)\,,\ X \sim \chi^2(n)\,,\ Y \perp X$$

- Tällöin satunnaismuuttuja t noudattaa **Studentin *t*-jakaumaa *n*:llä vapausasteella**.
- Merkintä:

$$t \sim t(n)$$

t-jakauman vapausasteet

- *t*-jakauman vapausasteiden lukumäärä *n* viittaa *yhteenlaskettavien lukumäärään* *t*-jakauman määrittelevän lausekkeen *nimittäjässä*.
- Vapausasteiden lukumäärä *n* on *t*-jakauman muodon määräävä *parametri*.

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $t \sim t(n)$.
- **Odotusarvo**:

$$\mathrm{E}(t) = 0,\ n > 1$$

- **Varianssi** ja **standardipoikkeama**:

$$\mathrm{Var}(t) = \mathrm{D}^2(t) = \frac{n}{n-2},\ n > 2$$

$$\mathrm{D}(t) = \sqrt{\frac{n}{n-2}},\ n > 2$$

Tiheysfunktion kuvaaja

- Kuva oikealla esittää *t*-jakauman

  $t(n)$

  tiheysfunktiota välillä [–4, +4], kun vapausasteiden lukumäärällä *n* on seuraavat arvot:

  (i) $n = 1$

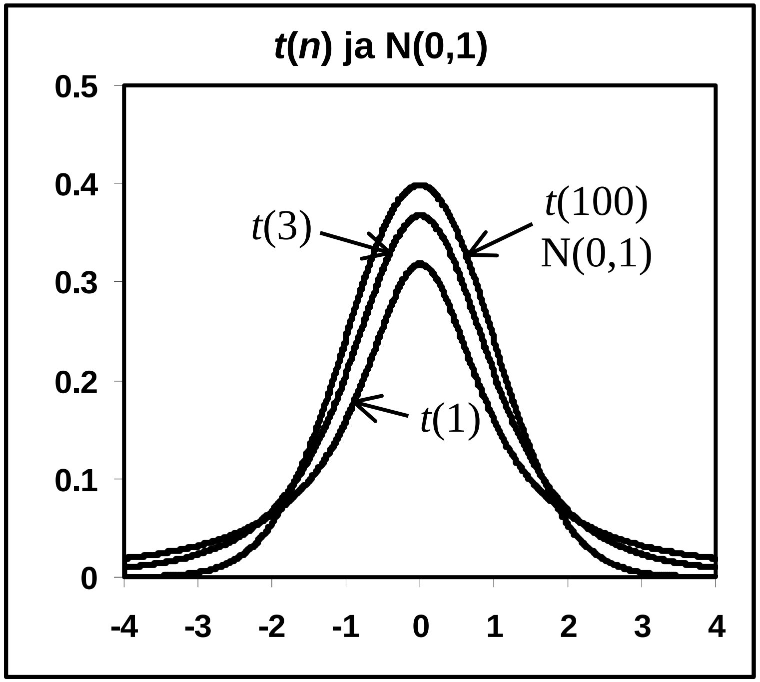
  (ii) $n = 3$

  (iii) $n = 100$

- Jakauman *odotusarvo*:

  $$\mathrm{E}(t) = 0 \,,\; n > 1$$

- Kuvaan on piirretty myös standardoidun normaalijakauman N(0,1) tiheysfunktion kuvaaja.



Tiheysfunktion ja sen kuvaajan ominaisuuksia 1/2

- *t*-jakauman tiheysfunktio $f(x)$ on kaikkialla *positiivinen*:

  $f(x) > 0$ kaikille x
- Tiheysfunktio on *yksihuippuinen*.
- Tiheysfunktio saa *maksimiarvonsa* pisteessä 0.
- Tiheysfunktio on *symmetrinen* suoran $x = 0$ suhteen:

  $f(-x) = f(+x)$ kaikille x

t-jakauma
Tiheysfunktion ja sen kuvaajan ominaisuuksia 2/2

- *t*-jakauman tiheysfunktio muistuttaa standardoidun normaalijakauman N(0,1) tiheysfunktiota, mutta on sitä *paksuhäntäisempi*.
- *t*-jakauman tiheysfunktio muistuttaa standardoidun normaalijakauman N(0,1) tiheysfunktiota *sitä voimakkaammin mitä suurempi on vapaus-asteiden lukumäärä n* (ks. tarkemmin >).

t-jakauma ja *F*-jakauma

- Olkoon $t \sim t(n)$.
- Tällöin

$$t^2 \sim F(1,n)$$

- Olkoon $F \sim F(1, n)$.
- Tällöin

$$\sqrt{F} \sim t(n)$$

t-jakauma ja normaalijakauma 1/2

- *t*-jakauma lähestyy standardoitua normaalijakaumaa, kun *vapausasteiden* lukumäärä *n* kasvaa.
- Olkoon $t \sim t(n)$.
- Tällöin

$$\lim_{n\to+\infty} \Pr(t \le z) = \Phi(z)$$

missä Φ on *standardoidun normaalijakauman* N(0,1) *kertymäfunktio*.

t-jakauma ja normaalijakauma 2/2

- Koska *t*-jakauma lähestyy vapausasteiden lukumäärän *n* kasvaessa standardoitua normaalijakaumaa N(0,1), voidaan *t*-jakaumaan liittyvät todennäköisyydet määrätä suurilla vapausasteiden luvuilla standardoidun normaalijakauman avulla.
- *Normaalijakauma-approksimaatio* *t*-jakaumalle on kohtuullinen jo, kun $n = 30$, ja riittävä useimpiin tarkoituksiin, kun $n > 100$.
- Esimerkki:

  Edellä esitetyssä kuvassa ei *t*(100)- ja N(0,1)-jakaumien tiheysfunktioiden kuvaajia pysty erottamaan toisistaan (ks. <).

Todennäköisyyksien määrääminen *t*-jakaumasta 1/2

- Todennäköisyyksien määrääminen *t*-jakaumasta voidaan tehdä jakauman *kertymäfunktion* avulla.
- Olkoon $t \sim t(n)$.
- Olkoon satunnaismuuttujan *t* *kertymäfunktio*

$$F_t(x\,;\,n) = \Pr(t \le x)$$

- Huomautus 1:

  Merkinnällä $F_t(x\,;\,n)$ on haluttu korostaa *t*-jakauman riippuvuutta sen vapausasteiden lukumäärästä n.

- Huomautus 2:

  t-jakauman tiheysfunktion integraalifunktiota ei tunneta, joten *t*-jakauman kertymäfunktion määräämiseen on käytettävä jotakin *numeerista menetelmää*.

Todennäköisyyksien määrääminen *t*-jakaumasta 2/2

- *Kaikkien* tapahtumien todennäköisyydet saadaan todennäköisyyksistä

$$\Pr(t \leq x) = F_t(x\,;\,n)$$

todennäköisyyslaskennan laskusääntöjen avulla.

- Esimerkiksi

$$\Pr(a \leq t \leq b) = F_t(b) - F_t(a)$$

Todennäköisyyksien määrääminen *t*-jakaumasta: Taulukot 1/3

- *t*-jakauman *taulukot* sisältävät tavallisesti *argumentin x arvoja* taulukoituna *useille vapausasteiden lukumäärille n*, mutta vain *muutamalle kertymäfunktion* F_t *arvolle*.
- Siten taulukot mahdollistavat seuraavan tehtävän ratkaisemisen (taulukkokohtaisin rajoituksin):

  Määrää *x*, kun todennäköisyys

  $$\Pr(t \le x) = F_t(x\,;\,n)$$

  on annettu.

Todennäköisyyksien määrääminen *t*-jakaumasta: Taulukot 2/3

- Koska *t*-jakaumaa käytetään tavallisesti *väliestimoinnin* tai *testauksen* yhteydessä, *t*-jakauman taulukoihin on yleensä taulukoitu sellaisia argumentin x arvoja, jotka vastaavat todennäköisyyden

  $$\Pr(t \le x) = F_t(x\,;\,n)$$

  komplementtitodennäköisyyttä

  $$p = \Pr(t \ge x) = 1 - F_t(x\,;\,n)$$

Todennäköisyyksien määrääminen *t*-jakaumasta: Taulukot 3/3

- Monissa *t*-jakauman taulukoissa on taulukoitu todennäköisyyksiä

  $$p = \Pr(t \geq x) = 1 - F_t(x\,;n)$$

  vain, kun $x \geq 0$.
- Tällöin todennäköisyydet $\Pr(t \leq -x)$ saadaan soveltamalla *t*-jakauman tiheysfunktion *symmetrisyyttä* pisteen $x = 0$ suhteen:

  $$\begin{aligned}\Pr\left(t \leq -x\right) &= 1 - \Pr(t \geq -x) \\ &= 1 - \Pr(t \leq x) \\ &= \Pr(t \geq x) \\ &= p\end{aligned}$$

Todennäköisyyksien määrääminen *t*-jakaumasta: Esimerkki

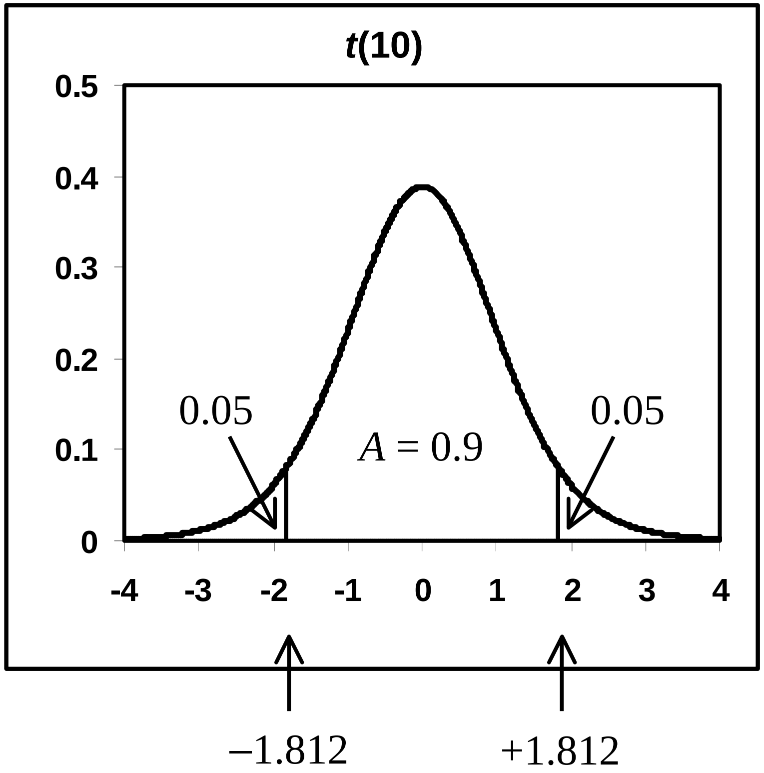
- Kuva oikealla esittää *t*-jakauman

  $t(10)$

  tiheysfunktiota välillä [0, 4].
- *t*-jakauman taulukoista saadaan:

  Alueen A pinta-ala

  $$= \Pr(-1.812 \le t \le +1.812)$$
  $$= F_t(+1.812;10) - F_t(-1.812;10)$$
  $$= 0.95 - 0.05$$
  $$= 0.9$$



Todennäköisyyksien määrääminen *t*-jakaumasta: Ohjelmat

- Olkoon $t \sim t(n)$.
- Monet *tietokoneohjelmat* mahdollistavat seuraavien tehtävien ratkaisemisen:

(i) Määrää todennäköisyys

$$\Pr(t \le x) = F_t(x\,;\,n)$$

kun x on annettu.

(ii) Määrää x, kun todennäköisyys

$$\Pr(t \le x) = F_t(x\,;\,n)$$

on annettu.