Laskuharjoitus 5A

10:43 » Toivottavasti kurssisivulla mainittu mainio, mutta ehkä liian haastava tehtävä saadaan jokatapauksessa mukaan esim. bonustehtävänä!
15:03 » Oispa jo ensi viikon laskarit ilmestyneet
16:04 » 15:03: nyt h5A on laitettu... – PetteriP (*)

12:42 » eikö h5at2:sessa lasketa p-arvoiksi vain yksittäisiä pistetodennäköisyyksiä? siis kun tilanne on esim, että 9 siementä oli itänyt, ei alle 9 eikä yli 9, vaan tasan. Tällöinhän p-arvo sille, että 9 siemenen itäessä theta=8/10 on P{K=9}, missä K kuvaa itäneiden siementen lukumäärää.
12:45 » jotenkin tuntuu hiukan hankalalta sisäistää tämä, mutta eihän se noin olekaan :D
12:45 » p-arvo: "Kyseessä on siis nollahypoteesin pätiessä laskettu yläraja sille todennäköisyydelle, tai yksinkertaisessa tapauksessa itse todennäköisyys, että satunnaismuuttuja T saa arvon, joka on yhtä suuri tai suurempi kuin nyt havaittu arvo t(y)." tämä ehkä selventää asiaa :)
13:09 » Itse yritän muistaa p-arvon tulkinnan yleensä muodossa "todennäköisyys sille, että jos nollahypoteesi olisi tosi, saataisiin arvoja, jotka ovat yhtä ääreviä tai äärevämpiä kuin havaittu aineisto". Tuo on minulle käsitteen idea: yritetään jotenkin kvantifioida sitä, kuinka "poikkeuksellisesta" tai "yllättävästä" tilanteesta on kyse. Teknisempi määrittely (esim. onko kyseessä yläraja, mitkä arvot ovat kriittisiä jne.) on sitten enemmän yksityiskohtien hiomista tilannetta vastaavaksi.
14:56 » h5at4, alpha=0.056 ?
18:08 » vinkkiä H5AT1? En vain oikein hoksaa mitä pitää tehdä :D Lähin ajatteluni on "jos tapahtumat ovat ekvivalentit, niin niillä on sama p-arvo", mutta en tiedä miten osoittaa tämä, osaan vain todeta :)
18:47 » 18:08: p-arvoa kannattanee lähteä tutkimaan määritelmän nojalla, mitä on määritelmän mukaan p-arvo testisuureelle t mielivaltaisella aineistolla y. Saat erään todennäköisyyslausekkeen, jossa todennäköisyyden P() sisällä on epäyhtälö. Sitten voit miettiä, voisiko se epäyhtälö olla ekvivalentti jonkin toisen epäyhtälön kanssa. Huomaathan, että mikä tahansa t:n transformaatio ei nähdäkseni johda ekvivalenttiin testiin. (Mitä, jos muunnos on t(y)^2, ja t(y) voi saada myös negatiivisia arvoja?)
18:50 » 18:47: Juuri tähän "P() sisällä on epäyhtälö. Sitten voit miettiä, voisiko se epäyhtälö olla ekvivalentti jonkin toisen epäyhtälön kanssa." olen päässyt, mutten ymmärrä miten siitä seuraa, että näiden kahden ekvivalentin epäyhtälön tn. eli p-arvo olisi sama? :)
18:54 » Tätä voinee ajatella pidemmän kaavan kautta tai sitten mutkia suoristaen. Muistat varmaan, miten käytettiin kertymäfunktiotekniikkaa: jos Y on sm ja Z = Y^3, niin F_Z(z) = P(Z <= z) = P(Y^3 < z) = P(Y < z^{1/3}) = F_Y(z^{1/3}). Idea on hyvin samantapainen: tn-mitan sisällä olevassa epäyhtälössä voi epäyhtälölle tehdä "normaaleja" operaatioita (=sellaisia, jotka säilyttävät järjestyksen) niin, että tn-mitta on edelleen sama.
18:57 » ...jos f on jokin kuvaus ja g on kuvaus x ⟼ x^3, niin mitä voit sanoa seuraavista epäyhtälöistä: "f(x) >= f(y)", "f(x)^3 >= f(y)^3" ja "g(f(x)) >= g(f(y))".
18:59 » Se, miten tässä vähän mutkia suoristettiin, on lähinnä merkinnällistä. P(X >= x):han on joukon {ω: X(ω) >= x} mitta, missä X on satunnaismuuttuja eli kuvaus. Koska X on kuvaus, niin tämä joukko on sama joukko kuin {ω: X(ω)^3 >= x^3}, joten sen tn-mittakin on sama.
19:03 » (Näemmä tiputin onnellisesti tuossa alla yhtäsuuruuden pois epäyhtälöstä kohdassa P(Z <= z) = P(Y^3 < z), mutta ehkä idea tuli lapsuksesta huolimatta selväksi. :))
19:55 » Jouduin käymään hetken pois. "kertymäfunktiotekniikka" -henkinen vinkki oli loistava, kiitos avusta erittäin paljon! :)
21:06 » 18:08,19:55: tuo mitä 18:47 ensin sanoi, että p-arvo testisuureelle s jonka suuret arvot ovat kriittisiä H_0:lle on (ainakin yksinkertaiselle H_0:lle) tn P(s(Y) ≥ s(y)). Jos nyt haluaa tarkastella

jotain s(Y):n muunnosta, niin tätä voi lähestyä kuten 18:54 oleellisestiTN2-kurssilta tutulla kertymäfunktiotekniikalla. :) – PetteriP (*)
21:08 » ... kun ajattelee p-arvoa "häntätodennäköisyytenä", niin lopulta siinä tarkastellaan testisuureen t(Y) jakaumaa, joten kaikki tn2:sen muunnoksiin liittyvät "kikkakolmoset" ovat mukavasti käytössä. – PetteriP (*)
21:10 » ... kun H_0 on yhdistetty, joutuu lopuksi vielä arvioimaan eri häntätodennäköisyyksien ylärajoja eli silloin on mukana hiukka lisää teknisiä haasteita. – PetteriP (*)
21:11 » 19:55: mutta olen samaa mieltä että "kertymäfunktiotekniikka" -henkinen vinkki oli loistava. Kiitos 18:54-19:55 :) – PetteriP (*)
21:21 » 12:42: kuten 12:45-13:09 sanoikin, niin tässä H_0 on "ainakin 80% itää", joten hyviä testisuureita olisi k(y) = "aineistossa y on k itänyttä" tai t(y) = "aineistossa y on t itämätöntä" = n - k(y) :) Nyt esimerkiksi pienet testisuureen k:n arvot ovat kriittisiä H_0:lle. Suuret arvot testisuureelle t ovat puolestaan kriittisiä... – PetteriP (*)
21:43 » ... siispä tuota t:tä voidaan käyttää kuten monisteessa, jolloin p = sup_{θ in Ω_0} P_θ (t(Y) ≥ t(y)). Käyttämässäsi 9 itänyt, vastaisi t(y) = "itämättä jäi" = 10 - 9 = 1, joten lasketaan tn:iä P_θ (T ≥ 1) = P_θ (10 - T ≦ 10 - 1) = P_θ (K ≦ 9). Eli kumpikin testisuure antaa kyllä samat p-arvot, kunhan tuon K:n p-arvoissa ajattelee pienet arvot kriittisiä tarkoittaa ≦ -merkkiä... – PetteriP (*)
21:43 » ... . Lisäksi K ~ Bin(10, θ) kun taas T ~ Bin(10, 1-θ) :) – PetteriP (*)
21:44 » 14:56: lasken huomenna aamuselta noita lukuarvoja :) – PetteriP (*)
23:26 » H5AT2, saan melko isoja p-arvoja, mahdankohan laskea ollenkaan oikein? Siis sellaisia, ettö ei olisi syytä hylätä väitettä. Mietin vaan, kun b-kohdan tehtävänanto viittaa siihen, että olisi...
08:56 » 23:26: oletan, että olet käyttänyt testisuureena k = "itäneitten lkm", jolloin p = sup P(k(Y) ≦ k(y)) ja kuten 21:43 mainitsin, kullakin henkilöllä K = k(Y) ~ Bin(10, θ), kun taas eri havaintoja vastaavat testisuureen arvot asiakas 1: k(y)= 9,... asiakas 4: k(y) = 8. Tarkastelen esimerkkinä kysymystä, jos A) jollakin asiakkaalla k(y) = 10 sekä B) jos jollakin k(y) = 1... – PetteriP (*)
09:02 » ... A) asiakkaan kaikki siemet itivät, joten p = sup P_θ (K ≦ 10). Mutta tiedämme, että binomijakautunut sm K voi saada arvoja {0,1,..,10}, joten tn P_θ (K ≦ 10) = 1 joten myös p = 1. Eli: "henkilö havaitsi kaikkien itävän, joten kukaan ei voi havaita parempaa". Tällöin A) ei olisi hylkäämässä H_0:aa millään merkitsevyystasolla α, sillä jos asiakas nyt hylkäisi valmistajan väitteen, voisi hylkäysvirheen tn olla jopa 1... – PetteriP (*)
09:10 » ... eli A) (jolla k(y) = 10) sai valtavan suuren p-arvon. Asiakas B) oli huono-onnisempi ja k(y) = 1. Nyt P_θ (K ≦ 1) = P_θ (K = 0) + P_θ (K = 1) = 1 × θ^0 × (1 - θ)^10 + 10 × θ^1 × (1 - θ)^9 = (1 - θ)^9 (10 θ + 1 - θ) = (1 - θ)^9 (9 θ + 1)... – PetteriP (*)
09:16 » ... eli asiakkaan B) p-arvo on p = sup_{θ ≥ 0.8} (1 - θ)^9 (9 θ + 1). Tämä sup on siis pienin alaraja (ja jos näiden tn:ien maksimi on olemassa, se on maksimi). Kuinkas tämä lasketaan? Jos kokeillaan arvoilla θ = 1, θ = 0.9 ja θ = 0.8 havaitaan, että vastaavat tn:t P_θ(K ≦ 1) = (1 - θ)^9 (9 θ + 1) ovat tn = 0, tn = (0.1)^9 × (8.1 + 1) = 9.1 × 10^{-9} ja tn = 0.2^9 × (7.2 + 1) = 8.2 × 2^9 × 10^{-9} = 4.1984 × 10^{-6}... – PetteriP (*)
09:25 » ... selvästi suurin arvo 4.2 × 10^{-6} = 0.0000042 saatiin kun θ = 0.8. Kysymys: miten valita θ siten, että tämä funktio saa suurimman arvon? Vastaus: tutkitaan derivaattaa :) Nyt ptnf P_θ(K = k) = c(k) × θ^k × (1-θ)^{10-k} vastaa uskottavuusfunktiota L(θ; k) ja siten log-uskottavuutta l(θ; k) = k log(θ) + (10-k) log(1-θ). Tämän derivaatta (eli pistemäärä) on l'(θ; k) = k/θ - (10-k) /(1-θ) = (k - 10×θ)/ (θ(1-θ)). – PetteriP (*)
09:29 » ... tämänhän olemme laskeneet jo esimerkissä 2.2.5 kun tutkimme ensimmäistä kertaa bernoullitoistokokeen (ja binomimallin) su-estimaattia. Huomaamme, että su-estimaatti on k/10 ja päättelimme tuolloin, että log-uskottavuusfunktio l on kasvava kun θ < k/10 ja vähenevä θ > k/10 (katsomalla derivaatan merkkiä :)... – PetteriP (*)
09:32 » ... ja jos log-uskottavuus on kasvava (vähenevä), niin uskottavuusfunktio L(θ; k) on vastaavasti kasvava (vähenevä). Siispä olemme päätelleet, että ptnf P_θ (K = k) on kasvava θ:n suhteen (vast. vähenevä), kun θ < k/10 (vast. kun θ > k/10)... – PetteriP (*)
09:37 » ... joten takaisin asiakkaan B) p-arvoon :) Nyt P_θ(K ≦ 1) = P_θ(K = 0) + P_θ(K = 1), joten edellisen perusteella P_θ(K = 0) on vähenevä, kun θ > 0 ja P_θ(K = 1) on vähenevä, kun θ > 0.1.

Siispä $P_\theta(K \leqq 1)$ on vähenevä, kun $\theta > 0.1$ joten välillä [0.8,1) se saavuttaa suurimman arvonsa kun $\theta = 0.8$. Siispä $p = \sup_{\theta \geq 0.1} P_\theta(K \leqq 1) = P_{0.8}(K \leqq 1) \approx 0.0000042$... – PetteriP (*)
09:38 » ... asiakkaan B p-arvo oli siis hurjan pieni ja tämä asiakas olisi valmis hylkäämään nollahypoteesin ei vain 5% vaan myös 0.001 merkitsevyystasolla. – PetteriP (*)
09:41 » 23:26: takaisin kysymykseesi. Osa p-arvoista on varsin isoja ja "tosi pieniä" (kuten huono-onnisen asiakkaan B tapauksessa) ei mukana ole. Mutta esimerkkini varmaankin antoi viitteitä kuka/ketkä saattaisivat olla hylkäämässä valmistajan väitettä 0.05:n merkitsevyystasolla... – PetteriP (*)
10:00 » En ole kysyjä, mutta kiitos tästä pitkästä asian avaamisesta. Hetkeksi tuli kyllä "kylmä hiki otsalle", sillä luulin aluksi, että tässä alla puheena ollut asiakas B olisi sen tehtävänannon asiakas, jonka siemenistä viisi on itänyt. Eli ihan noin pientä p-arvoa ei hänelle tehtävässä tule, mutta tuon esimerkin B:lle tulee tietty :)
10:03 » 14:56: aivan kuten laskitkin, niin merkitsevyystaso alpha on tuo ja tuo luku on tarkkaan esitettävissä kahden pienen alkuluvun potenssien osamääränä :) – PetteriP (*)
10:04 » H5AT5:seen liittyen: en oikein keksinyt tähän parempaa lähestymistapaa kuin keskeinen raja-arvolause -> eräs muunnoksia käsittelevä lause tn2 kurssilta -> valmis. Menikōhän liian helposti ja jotain oikomalla. Tosin itse en mielestäni "aukkoja" omasta päättelystä löytänyt :) ja ps. tännehän onkin tullut paljon teksiä eilen ja tänään, mikä on aina kiva juttu :)
10:08 » 10:00: aivan :) eihän sitä nyt sovi tehtävää täysin "spoilata", mutta ajattelin, että tämä on hyvä kohta aukaista se, miten "pienet arvot kriittisiä" / "suuret arvot kriittisiä" liittyy mihinkään sekä varsinkin, että p-arvon laskeminen on 1) häntätn-askemista testisuureelle ja 2) yhdistetyn H_0:n tapauksessa vielä ylärajan hakemista... – PetteriP (*)
10:08 » ... ja 3) kuinka uskottavuuspäättelymme itse asiassa kertoo vallan suoraan, millä arvolla θ tuo maksimi saavutetaan. Jos olisikin että $H_0 : \theta > 0.8$, niin varsinaisesti maksimia ei saavuteta, mutta "suppi" (puhekielen supremum eli pienin yläraja) saavutetaan silti kun $\theta = 0.8$:) – PetteriP (*)
10:12 » 10:04: keskeinen raja-arvolause onkin vallan mainio työkalu. Uskoisin, että meni juuri oikein :) Tarkoitukseni on korostaa, että estimaattorien avulla voidaan rakentaa testisuureita ja että normaaliapproksimaatiolla saadaan mukavast "rotia" jakaumatarkasteluihin, joita tarvitaan p-arvo / kriittinen alue / voimafunktio -pohdintoihin :) – PetteriP (*)
12:35 » kiitos vastauksesta 10:12. Uskoisin, että kärryillä siis ollaan :) Aluksi tehtävä vaikutti tosi hankalalta, mutta näiden muutamien mukavien apulauseiden avulla päästään melko vaivatta näkemään mikä (approksimoitu) jakauma on kyseessä.