Kertaus ja kurssikoe

18:34 » Tarvitseeko tenttiin erikseen ilmoittautua, kun se on tenttiviikon jälkeisellä viikolla?
18:58 » 18:34: ei tarvitse, kyseessä on normaali kurssikoe vaikka sen aika onkin hieman normaalista poikkeava. – PetteriP (*)

05:54 » Jos joku käy tänään kertausluennolla ja kirjoittaa muistiinpanoja niin arvostaisin jos voisi laittaa tänne dropboxin/muun tiedostonjakelupalvelun kautta muistiinpanot. Kiitoksia etukäteen :)
18:46 » Korvaavasta kokeesta mitään uutta tietoa? Tai pystyykö kurssin tenttimään yleistentissä huhtikuussa?
21:16 » 18:46: yritän huomenna aamupäivästä selvittää varmistaa tuon, mutta uskoisin että huhtikuu on hyvinkin mahdollinen. Mutta varmistan tuon siis vielä. – PetteriP (*)
20:27 » Onko tulossa ratkaisuehdotus 1.2. pidettyyn TN II:n 2. kurssikokeen korvaavaan kokeeseen?
21:17 » 20:27: tulee kyllä mutta niissä menee vielä "tovi". – PetteriP (*)
10:10 » Mitä kaikkia tietoja tentissä annetaan? Varmastikin ainakin jokin taulukko, jota tarvitsee luottamusväleihin, mutta annetaanko muutakin? Lähinnä pohdin että täytyykö esim. kaikkien mahdollisten jakaumien tiheysfunktiot yms. kirjoittaa lunttiin?
10:44 » 10:10: Jakaumista annetaan tuollaiset tarvittavat perustiedot.
10:45 » ..."kaikkien mahdollisten" jakaumien tiheysfunktioiden kirjoittaminen olisikin aika villi tehtävä! ;D
11:54 » 10:10: tentissä annan tehtäväpaperin ohessa tarvittavat taulukot ja perustiedot tarvittavista jakaumat. Ja tuo "kaikkien mahdollisten" kirjoittaminen olisi aika haastava tehtävä aivan kuten 10:45 mainitsi. Päivitän tietoja kurssisivulle myöhemmin tänään (pahoittelut hiljaisuudesta :) – PetteriP (*)
18:17 » Miten on tuon korvaavan kurssikokeen laita? Miten siihen ilmottaudutaan ja milloin (:
08:15 » 18:17: korvaava kurssikoe on yleistentin yhteydessä 12.4. ja ilmoittautuminen s-postilla. Laitoin tietoja hieman jo kurssisivulle ja päivitän sitä myös :) – PetteriP (*)

15:48 » Olen kertaamassa kokeeseen ja törmäsin seuraavaan ongelmaan: Oletetaan, että Y_1,...,Y_n ovat riippumattomia ja samoin jakautuneita. Miten E(n/(ΣY_i)) kannattaa laskea, kun nyt ei voida käyttää odotusarvon lineaarisuutta?
16:30 » Onkohan harjoituksen 5A tehtävän 2 ratkaisussa nyt ihan oikein laskettu nuo a-kohdan p-arvot. Todella isoja p-arvoja ja tolla annettulla kaavalla saisi p-arvoksi tasan 1, jos havainto olisi k=10. Eikö myös havainnolla k=8 pitäis olla isompi p-arvo kuin havainnolla k=9, kun nollahypoteesina on käytetty theta = 0.8.
19:07 » Kysymys harjoituksen 5B tehtävästä 1: Englanninkielisen Wikipedian mukaan Fisherin informaatio olisi 1/(p(1-p)), kun taas ratkaisuehdotuksessa on saatu n/(p(1-p)). Onko Wikipedia väärässä?
19:58 » Wikipedian tulos tarkoittanee yhden havainnon mallia. – Aku
21:12 » 16:30: monisteessa p-arvo määriteltiin tunnusluvulle T, jonka suuret arvot ovat kriittisiä nollahypoteesille H_0: θ ≥ 0.8. Eräs tällainen voisi siis olla t = 10 - k, sillä jos t on suuri, niin k on pieni. Laskemme siis P_θ (t(Y) ≥ t(y)) = P_θ (10 - K ≥ 10 - k) = P_θ (K ≦ k). Jos k = 10, niin oli theta mikä tahansa, niin P_θ (K ≦ 10) = 1, joten aivan kuin totesit p-arvo olisi tasan 1... – PetteriP (*)
21:12 » ... Eräs p-arvon tulkinta on "yläraja tn:lle, että saisimme tätä korkeintaan yhtä hyvin sopusoinnussa olevia havaintoja". Mutta emmehän me voi saada havaintoja joilla k > 10, joten tämä

p-arvo = 1 on yhteensopiva tämän kanssa :) Tehtävässä esiintynyt k = 9 tapauksessa P_θ (K ≦ 9) = 1 - P_θ (K = 10) = 1 - θ^{10}... – PetteriP (*)
21:13 » ... Näemme derivoimalla, että tämä on vähenevä θ:n suhteen joten p = sup_{θ ≥ 0.8} P_θ (K ≦ 9) = P_{0.8} (K ≦ 9) = 1 - 0.8^{10} = 0.89262... Eli p-arvo on todellakin valtava myös tässä tapauksessa. Eli tosiaankin: ainoastaan yksi havainnoista antaa pienehkön p-arvon... – PetteriP (*)
21:14 » ... Lisäksi kun havaintona on k = 8, niin huomaamme, että P_θ (K ≦ 8) < P_θ (K ≦ 9) ≦ 0.89262. Voimme siis päätellä, että p-arvon kun k=8 on oltava korkeintaan 0.89262 ja voisimme osoittaa, että se saa maksiminsa kun θ = 0.8 ja arvo on sama kuin mitä ratkaisuehdotuksessa. Selvensiköhän tämä kysymystäsi? – PetteriP (*)
21:37 » 15:48: hmm... noilla tiedoilla en vielä osaa sanoa sopivinta tapaa :) Mutta voisit yrittää seuraavaa. Jos tiedämme, että Y_1 + ... + Y_n = S noudattaa tunnettua jakaumaa, niin silloin voisimme kokeilla TTL:ää, sillä E(n/sum_i Y_i) = E(n/S) = Eg(S), kun g(x) = n/x... – PetteriP (*)
21:39 » ... 2) jos emme tunnista X:n jakaumaa, niin voisimme laskea n-kertaisen konvoluution :D (eli ei ehkä suositeltava), mutta 3) jos olettaisimme, että Y_i:llä on äärellinen varianssi σ^2 ja odotusarvo μ^2 ja n on valtavan suuri, niin tällöin asymptoottisesti S ~ N(nμ, nσ^2) ja asymptoottisesti X_n = S/n ~ N(μ, σ^2/n)... – PetteriP (*)
21:56 » ... ja pienellä "pulijauksella" (ja hieman tarkemmin käyttämällä TN2-kurssin luvun 11.5 lausetta 11.4 (eli delta-menetelmää) osoittaa, että itse asiassa (kunhan μ ei ole nolla) 1/X_n = n/(sum_i Y_i) ~ N(1/μ , σ^2/(μ^2 * n)) asymptoottisesti, joten E(1/X_n) ≈ 1/μ suurilla n. – PetteriP (*)
22:36 » 21:37-21:56: Y_i ~ Geom(p) ja tuo n/(ΣY_i) on laskemani SU-estimaattori, jonka odotusarvon yritän selvittää. Kiitos viesteistä! Alan pohtia :)
11:14 » 22:36: mainiota :) Tuossa voit kyllä laskea summan ΣY_i jakauman tarkkaan (vihje: negatiivinen binomijakauma, huom: TN2:sen geometrinen jakauma ja TP2:n geometrinen jakauma eivät ole ihan samat (ero Y= X + 1)). Mutta kokeile laskea odotusarvo ensin kun n = 1 (ja jos haluat vihjeen vastauksesta kun n = 1, niin kerro :). – PetteriP (*)
20:45 » 11:14: En ole alkuperäinen kysyjä, mutta kysymys kuulosti mielenkiintoiselta ja nyt olisin vinkin tarpeessa. :) Jos X = sum Y_i, niin sain parillakin eri tapaa pääteltyä (joko ajattelemalla sarjaa Bernoulli-kokeita tai sitten geometrisen jakauman mef:n avulla), että X = Z + n, missä Z ~ Negbin(n, p). Menikö tähän asti oikein?
20:45 » (Tässä geometrinen jakauma on ymmärretty TP2:n sopimuksin, negatiivinen binomijakauma sen sijaan TN2:n sopimuksin, eli Z = "epäonnistumisten lukumäärä ennen n:ttä onnistumista".)
20:48 » Sitten yritin laskea odotusarvoa sm:lle n/(Z+n) estimaattorille TTL:llä ja muunnosfunktiolla g(z) = n/(z+n). Tästä sain sarjan sum n/(z+n) * negbin(n, p)-jakauman tf. Keksin lähinnä triviaalin tavan (mielestäni) osoittaa, että se suppenee (koska osamäärä kasvaa pienentämällä nimittäjää ja z on ei-neg. luku, valitaan majorantiksi sum n/n * negbin(n, p) = 1), mutta ihan valtavan hyödylliseltä johtopäätökseltä tämä ei tuntunut. :)
20:49 » ...eli pitäisikö sarja saada "pyöriteltyä" sopivampaan muotoon, olisiko syytä käyttää alun perinkin jotain toisenlaista tapaa mallintaa se, vai tuliko laskuissa jossain kohtaa päättely-/laskuvirhe?
01:04 » Olikos niin, että jos ei aivan täysin ole sisäistänyt monisteen/kalvojen osioita 6.3-6.4 (approksimatiiviset uskottavuusosamäärään ja Waldin testiin perustuvat luottamusjoukot), ei tarvitse peljästyä? Niistähän ei tainnut olla hirveästi harjoituksia tai (muistini mukaan) luennoillakaan kovin yksityiskohtaisesti puhetta. :)
11:09 » 20:45-20:49: olet täysin kärryillä ja sarja on mahdollista saada pyöriteltyä sopivaan muotoon... Kirjoitan siitä hieman myöhemmin tänään :) – PetteriP (*)
11:17 » 01:04: ei kauheasti tarvitse huolestua. Kappaletta 6.4 emme juuri ennättäneet käydä, kappeleen 6.3. kohdista kävimme vielä tiistaina harjoitustehtävien ulkopuolelta alaluvut 6.3.1 ja 6.3.2. (luottamusväli <-> uskottavuusväli). Puhuin myös profiiliuskottavuudesta, mutta: "näistä ei siis ollut harjoituksia, joten ei tarvitse peljästyä :)" – PetteriP (*)

16:30 » Onko tentin kannalta keskeistä osata osoittaa jokin malli säännölliseksi? Kaikki harjoituksissa käytetyt mallit on taidettu vain todeta tarpeeksi säännöllisiksi...
17:16 » 16:30: Petteri varmaan vastaa tähän tarvittaessa tarkemmin, mutta "ensiapuna" uskaltaisin tyynnyttää mieltä, että ei tarvitse. :) Kaikkien säännöllisyysehtojen osoittaminen ei taida ilman "lisäkoneistoa" edes onnistua tämän kurssin puitteissa. Riittänee muistaa sen tyyppisiä tietoja kuin että a) eksponenttiperheen jakaumat ovat säännöllisiä ja b) esim. tasajakauma Tas(0, theta) ei ole (kantaja riippuu parametrista).
22:02 » 16:30: ei ole tarpeen osoittaa :) Eli kun käsitellään "tuttuja" jakaumia, niin näistä vain "havainnot tasajakaumasta" -malli ei ole säännöllinen. Eli sen saa vapaasti olettaa olevan voimassa (ja kerron kyllä ellei se ole (poislukien tuo tasajakauma :)) – PetteriP (*)
22:03 » ... Olen tehnyt kertaustehtävien ratkaisuehdotuksia, mutta skannausongelman takia saan ne kurssisivulle vasta huomenna. Pahoitteluni. – PetteriP (*)
11:31 » Kurssisivuilla oli maininta, että koepaperiin tulee "perustiedot tarvittavista jakaumista". Voiko olettaa, että nämä olisivat nuo kurssimonisteen liitteessä (jakaumia) olevat tiedot? :)
14:03 » Kertaustehtävä 1/8 eli monisteen tehtävä 3.8: tehtävänä oli osoittaa, että jakaumalla Bin(n, theta) parametrille 1/theta ei ole harhatonta estimaattoria olemassa. Eikö tämä seuraa suoraan siitä, että diskreetin sm:n odotusarvo on määritelmän mukaan olemassa vain, jos summa |t(k) f_K(k)| suppenee eli osasummien jono on ylhäältä rajoitettu, mutta kun theta -> 0+, niin 1/theta -> +inf (eli mitään ylärajaa nimenomaan ei ole olemassa)?

14:40 » Tietääkö joku milloin kesäkurssien tiedot/ennakkotiedot tulee?
14:52 » 14:40: Viime vuonna päätökset tulivat ~huhtikuun puolessavälissä. Kesäopetuskysely tuli juurikin näillä main kevättä, ja sen mukana ennakkotiedot kursseista, jotka "joka tapauksessa järjestetään". Avoimen kesäopinnot löytyvät jo täältä: https://www.helsinki.fi/fi/uutiset/kesaopinnot-2017

16:05 » 11:31: kyllä voi olettaa + lisätietoja, mitä siellä ei ole (tietoja TN2-kurssilta löytyviä, jos niitä tarvitaan :) Kuten yhteenlaskuominaisuuksia (tyylilin: jos X ~ P(μ), Y ~ P(λ), X ja Y riippumattomia, niin X + Y ~ P(μ + λ) jne. :) – PetteriP (*)
16:12 » 14:03: kyllä tuo riittää :) eli jos T on mikä vain estimaattori, niin väistämättä E_θ T < M jollakin M riippumatta θ:sta. Ja siispä 1/θ pienillä θ on väistämättä aidosti > kuin E_θ T. – PetteriP (*)
21:58 » Kysyisin vähän merkinnöistä: Kirjallisuudessa $(\Sigma x^2)/n$ merkitään joskus siten, että x^2:n päällä on viiva (eli viiva on sekä x:n että ²:n päällä). Saako tällaisia merkintöjä käyttää tentissä?
22:06 » 21:58: Itse olen yleensä erikseen määritellyt käyttämäni apumuuttujat (esim. logaritmien keskiarvo, neliöiden keskiarvo jne.), jos ne eivät ole olleet aivan ilmeisiä. Silloin tohtisin väittää niiden käytön olevan melko varmasti sallittua (ja usein helpottavan sekä kirjoittajan että lukijan hommaa). :)
12:55 » 21:58: kirjallisuudessa käytetään myös <x^2> merkintöjä (ja muita :) Kokeessa voit käyttää erilaisia merkintöjä, jos luotat siihen, että ymmärrän mitä tarkoitat :) Eli mitä ilmeisempi, sen parempi :) Jos monisteessa käytetty, varmasti ok. Kurssilla on käytetty yläviivaa kuvaamaan otoskeskiarvoa. Jos nyt vaikka kirjoittaa, että "olkoon jatkossa w_i = x_i^2" niin silloin taatusti ymmärrän merkinnän "w ja viiva päällä" tarkoittavan samaa kuin (Σx_i^2)/n. :) – PetteriP (*)
13:42 » Onko niin, että uskottavuusosamäärän testisuureelle ei päde asymptoottinen jakaumatulos, että r^(1/2) noudattaisi jakaumaa N(0, 1)? (Vrt. Waldin ja Raon testisuureille?) Jos, niin miksi ei? :)
14:52 » Kysymys koskien kert.tehtävien mallivastauksia (kertaus1-osa2.pdf): ensimmäisessä tehtävässä tilastollisen mallin nimittäjässä on y_i ja potenssissa joku pallero, mikä se on? Kakkostehtävän puolessavälissä ihmetyttää nimittäjissä esiintyvät nollat.
15:01 » 14:52: Huutomerkeistä (leveistä sellaisista) ensimmäisessä ja thetoista toisessa tapauksessa lienee kyse, jos ymmärrän oikein, mihin symboleihin viittaat. :)
15:04 » No niinpä tietysti! Kiitos!

16:01 » 13:42: periaate on suunnilleen se, että $r(Y)$ on likimain suurilla sama kuin $w(Y)$ (eli Waldin testisuure). Jos otamme siitä neliöjuuren, saamme likimain $|w(Y)|$:n eli sen asymptoottien jakauma on kutamain halfnormal distribution (kääntäisin suomeksi puolinormaalijakauma). Eli X:stä pääsee helposti X^2:een mutta X^2:sta pääsee $|X|$:ään :) – PetteriP (*)
16:07 » Ah niin, arvelinkin että ongelma jotenkin liittyisi juuren ottamiseen neliöstä, mutta en sitten ihan loppuun asti ajatellut. :) Kiitos!
18:08 » Kiitoksia erilliskokeen ratkaisuehdotuksista! Niistä taisi unohtua(?) kirjoittaa tai skannailla viimeisen tehtävän eli 4b-kohdan vastaus eli testaaminen testisuureilla. (Vahinko ei liene kuitenkaan kovin suuri, koska p-arvojen laskeminen on aika mekaanista, kun testisuureet on laskettu, ja näistähän löytyy myös harjoituksista esimerkkejä. :)
19:58 » 18:08: no jopas... skannerin ohisyöttö imaisee välillä pari sivua kerralla, joten .... Mutta nopeasti: uskottavuusosam. testin $p = P_{1000}(r(Y) \geq r(y)) \approx P(Z \geq 2.314)$ ja Waldille sekä Raolle $p = P_{1000}(u(Y) \geq u(y)) \approx P(Z \geq 2)$, missä Z noudattaa khiin neliön jakaumaa vapausasteella 1 (kaikissa suuret arvot kriittisiä H_0:lle). – PetteriP (*)
19:58 » ... Taulukosta näemme, että $0.10 = P(Z \geq 2.706) < P(Z \geq 2.314) < P(Z \geq 2)$, joten kukin p-arvoista $> 0.10 > 0.05$ mikä oli tehtävän merkitsevyystaso. Eli nollahypoteesi H_0 on syytä jättää voimaan :) – PetteriP (*)