TILASTOMATEMATIIKKA

Harjoitus 3 ratkaisut, kevät 2018

Harjoituksen teemoja ovat

(i) Diskreetin jakauman odotusarvo, varianssi ja keskihajonta

(ii) Jatkuvan jakauman odotusarvo, varianssi ja keskihajonta

(iii) Odotusarvon ja varianssin ominaisuudet

(iv) Keskeinen raja-arvolause

(v) Normaalijakauma-approksimaatio

(vi) Jatkuvuuskorjaus

Tehtävät 1-5 liittyvät teemoihin (i)-(iii), joita käsitellään pääasiassa alkuviikon harjoituksissa. Teemoja (iv)-(vi) käsitellään tehtävissä 6-9, jotka on ensisijaisesti tarkoitettu loppuviikon tehtäviksi. Tehtäviä voi toki ratkoa haluamassaan järjestyksessä. **Minimivaatimuksena laskuharjoituspisteille** ovat **tehtävät 2 ja 3** alkuviikosta ja **tehtävät 6 ja 7** loppuviikosta.

1. Olkoon X kruunujen lukumäärää kolmen kolikon heitossa.

   a) Määrää X:n odotusarvo ja varianssi X:n pistetodennäköisyyksien avulla.

   b) Kohdan a) tapa ei ole kovin hedelmällinen, kun kolikkoja on useita. Sen vuoksi kannattaa miettiä mikä on X:n jakauma. Määrää odotusarvo ja varianssi jakaumamallin avulla.

   c) Mitä on kruunujen lukumäärän odotusarvo ja varianssi, kun kolikkoja on sata?

Ratkaisu:

a) Alkeistapahtumat kolmen kolikon heitossa voidaan luetella:
$S = \{$HHH, HHT, HTH, THH, HTT, THT, TTH, TTT $\}$, missä H = kruunu(Head) ja T = klaava(Tail). Pistetodennäköisyydet kruunujen lukumäärälle X ovat siis

$$P(X=0) = P(X=3) = \frac{1}{8}$$
$$P(X=1) = P(X=2) = \frac{3}{8}$$

Tällöin odotusarvo on

$$E(X) = \sum_{k=0}^{3} kP(X=k) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \underline{\frac{3}{2}}$$

Varianssia varten lasketaan ensin

$$E(X^2) = \sum_{k=0}^{3} k^2 P(X=k) = 0^2 \cdot \frac{1}{8} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8} = 3$$

Tästä saadaan

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = \underline{\frac{3}{4}}$$

b) Kolmen kolikon heittoa voidaan pitää toistokokeena ja silloin kruunujen lukumäärä $X \sim \mathrm{Bin}(n, p)$, missä $n = 3$ ja $p = \frac{1}{2}$. Odotusarvo on tällöin

$$E(X) = np = 3 \cdot \frac{1}{2} = \underline{\frac{3}{2}}$$

ja varianssi on

$$\mathrm{Var}(X) = np(1-p) = 3 \cdot \frac{1}{2}(1 - \frac{1}{2}) = \underline{\frac{3}{4}}$$

c) Nyt $n = 100$, joten

$$E(X) = np = 100 \cdot \frac{1}{2} = \underline{50}$$

$$\mathrm{Var}(X) = np(1-p) = 100 \cdot \frac{1}{2}(1 - \frac{1}{2}) = \underline{25}$$

2. Teekkareiden laskiaisriehassa järjestetään arpajaiset, jossa on myynnissä 1000 arpaa. Voittoja jaetaan seuraavasti: päävoitto on 100 euroa, 10 kappaletta 10 euron voittoja ja 50 kappaletta 5 euron voittoja.

   a) Turo ostaa yhden arvan. Määrää Turon voittotodennäköisyysjakauma.

   b) Mikä on odotettavissa olevan voiton suuruus ja hajonta?

   c) Mikä arvan hinnan pitää olla, että arpajaiset kannattavat, kun arpoja ostetaan puolet?

Ratkaisu: Olkoon X = ”Turon yhdellä arvalla saama voittosumma”.

a) Muuttujan X arvojoukko on $S_X = \{0, 5, 10, 100\}$. Päävoiton voittotodennäköisyys on $\frac{1}{1000}$, 10 euron voiton todennäköisyys on $\frac{1}{100}$,... Jakauma voidaan ilmoittaa vaikkapa taulukkona

Voittosumma	0	5	10	100
Voittotodennäköisyys	$\frac{939}{1000}$	$\frac{1}{20}$	$\frac{1}{100}$	$\frac{1}{1000}$

b) Voiton odotusarvo on

$$\mu = E(X) = 0 \cdot P(X = 0) + 5 \cdot P(X = 5) + 10 \cdot P(X = 10) + 100 \cdot P(X = 100) = \frac{45}{100}.$$

Varianssiksi saadaan

$$\begin{aligned}\mathrm{Var}(X) =& P(X = 0)(0 - \mu)^2 + P(X = 5)(5 - \mu)^2 + P(X = 10)(10 - \mu)^2 \\ &+ P(X = 100)(100 - \mu)^2.\end{aligned}$$

Sijoittamalla laskettu odotusarvo paikalleen ja ottamalla neliöjuuri saadaan hajonnaksi

$$\sigma_X \approx 3.5.$$

Voiton odotusarvo on siis 45 senttiä ja hajonta 3.5 €.

c) Oletetaan, että kaikki voittoarvat myydään, jolloin voittoja jaetaan yhteensä 450 €:n edestä. Koska arpoja ostetaan puolet, saadaan yhden arvan hinnaksi 90 senttiä, jotta menot voidaan kattaa.

3. Suurissa kaupungeissa on tunnetusti paljon liikennevaloja. Odotusaika liikennevaloissa on satunnaismuuttuja, joka vaihtelee 0 ja 2 minuutin välillä. Odotusajan tiheysfunktio saa maksimiarvonsa 0:ssa ja pienenee lineaarisesti nollaan 2 minuutin kohdalla.

   a) Määrää tiheysfunktion lauseke, keskimääräinen odotusaika ja odotusajan varianssi.

   b) Kalle Kiihdyttäjä tuntee kaupungin poikki reitin, jonka varrella on vain 10 liikennevaloa. Laske koko reitin yhteisen odotusajan keskimääräinen arvo ja keskihajonta.

Ratkaisu:
Merkitään odotusaikaa liikennevaloissa satunnaismuuttujalla X.

a) X:n jakauman tiheysfunktio on

$$f(t) = \begin{cases} 1 - \frac{1}{2}t & \text{kun } 0 < t < 2 \\ 0 & \text{muulloin} \end{cases}$$

$$E(X) = \int_{-\infty}^{\infty} tf(t)dt = \int_0^2 t(1 - \frac{1}{2}t)dt = \Big/_0^2 (\frac{1}{2}t^2 - \frac{1}{2} \cdot \frac{1}{3}t^3) = \underline{\frac{2}{3}}$$

$$\begin{aligned} \mathrm{Var}(X) &= \int_{-\infty}^{\infty} (t - E(X))^2 f(t)dt \\ &= \int_0^2 (t - \frac{2}{3})^2 (1 - \frac{1}{2}t)dt \\ &= \int_0^2 (-\frac{1}{2}t^3 + \frac{5}{3}t^2 - \frac{14}{9}t + \frac{4}{9}) \\ &= \Big/_0^2 (-\frac{1}{8}t^4 + \frac{5}{9}t^3 - \frac{7}{9}t^2 + \frac{4}{9}t) \\ &= \underline{\frac{2}{9}} \end{aligned}$$

b) Olkoon odotusaika 10 liikennevalossa T ja kussakin i:nnessä valossa odotusaika X_i, jolloin $T = \sum_{i=1}^{10} X_i$.

$$\begin{aligned} E(T) &= E(X_1 + X_2 + \cdots + X_{10}) \\ &= E(X_1) + E(X_2) + \cdots + E(X_{10}) = 10 \cdot \frac{2}{3} = \underline{\frac{20}{3}} \\ \mathrm{Var}(T) &= \mathrm{Var}(X_1 + X_2 + \cdots + X_{10}) \\ &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_{10}) = 10 \cdot \frac{2}{9} = \frac{20}{9} \end{aligned}$$

T:n keskihajonta on siis $\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{20}{9}} = \underline{\frac{2\sqrt{5}}{3}}$.

4. Magneettisessa tiedon tallennuksessa oli havaittu virheiden määrän suurta talletusyksikköä kohti noudattavan Poisson-jakaumaa siten, että talletuksessa esiintyi keskimäärin yksi virheellinen bitti 100000 talletettua bittiä kohti. Tieto talletetaan sektoreina, missä sektori käsitti 4096 8-bittistä sanaa.

   a) Kuinka monta tallennusvirhettä esiintyi keskimäärin yhdessä sektorissa?

   b) Kuinka monta sektoria on keskimäärin käytävä läpi jotta ainakin yksi tallennusvirhe löytyy? Sektorit ovat toisistaan riippumattomat.

Ratkaisu:
Meitä kiinnostava satunnaismuuttuja on $X =$ ”virheiden lukumäärä sektoria kohden”, joka on tehtävänannon mukaan Poisson-jakautunut, $X \sim \text{Poi}(a)$ parametrilla

$$a = \frac{4096 \cdot 8}{100000} = \frac{32768}{100000} \approx 0.328,$$

sillä yksi sektori pitää sisällään 32768 bittiä.

a) Koska Poisson-jakauman parametri a ilmoittaa keskimääräisen arvon, eli $E(X) = a$, niin yhdessä sektorissa esiintyy keskimäärin

$$E(X) = a \approx 0.328$$

tallennusvirhettä.

b) Olkoon Y vähintään yhden tallennusvirheen sisältävien sektoreiden lukumäärä, jolloin $Y \sim \text{Bin}(n, p)$, missä n on tuntematon ja p on todennäköisyys sille, että sektorissa on vähintään yksi tallennusvirhe. Tehtävänannon mukaan

$$p = P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{a^0}{0!}e^{-a} = 1 - e^{-\frac{32768}{100000}} \approx 0.279.$$

Tällöin ainakin yhden tallennusvirheen sisältäviä sektoreita on keskimäärin

$$E(Y) = np$$

kappaletta. Koska tallennusvirheitä pitää olla keskimäärin ainakin yksi, saadaan

$$E(Y) = np \geq 1 \Rightarrow n \geq \frac{1}{p} \approx 3.58,$$

eli on käytävä läpi keskimäärin 4 sektoria vähintään yhden virheen löytämiseksi.
Tätä voidaan ajatella myös geometrisen jakauman, joka löytyy luentomonisteen sivulta 30. Olkoon Y sen sektorin järjestysluku, jolla esiintyy ensimmäinen tallennusvirhe. Tällöin Y noudattaa geometrista jakaumaa parametrilla p ja merkitään $Y \sim \text{Geo}(p)$, missä p on todennäköisyys sille, että sektorissa on vähintään yksi virhe, eli $p = P(X \geq 1)$. Luentomonisteen sivun 44 mukaan

$$E(Y) = \frac{1}{p} \approx 3.58.$$

5. Erään kemikaalin pitoisuus $[g/cm^3]$ tietyssä aineseoksessa on normaalijakautunut satunnaismuuttuja, jolla on odotusarvo μ ja hajonta $\sigma = 0.004\ g/cm^3$. Tehtävänä on arvioida odotusarvoa μ otoksen avulla. Kuinka monta toisistaan riippumatonta näytettä on otettava, jotta näytteestä mitattujen kemikaalipitoisuuksien keskiarvo poikkeaisi μ:stä suuntaan tai toiseen korkeintaan $0.002\ g/cm^3$ todennäköisyydellä 0,9?

Ratkaisu:
Olkoon $X_i =$ "kemikaalin pitoisuus näytteessä i". Tehtävänannon mukaan $X \sim N(\mu, \sigma^2)$. Kemikaalipitoisuuden keskiarvoa kuvaava satunnaismuuttuja on aritmeettinen keskiarvo

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Luentokalvojen Esimerkin 23 ja sivun 41 mukaan $\overline{X}$ on normaalijakautunut parametreilla

$$E(\overline{X}) = \mu \quad \text{ja} \quad \sigma^2_{\overline{X}} = \frac{\sigma^2}{n}.$$

Tähän palataan useasti myöhemminkin.
Yllä näytteiden lukumäärä n on tuntematon suure, joka ratkaistaan ehdosta

$$P(\mu - 0.002 \leq \overline{X} \leq \mu + 0.002) \geq 0.9.$$

Standardisoidaan $\overline{X}$, jolloin

$$Z = \frac{\overline{X} - E(\overline{X})}{\sigma_{\overline{X}}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

ja päästään taulukkoarvoihin käsiksi. Standardisoimalla saadaan

$$\begin{aligned}
P(\mu - 0.002 \leq \overline{X} \leq \mu + 0.002) &= P(\mu + 0.002 - \mu \leq \overline{X} - \mu \leq \mu + 0.002 - \mu) \\
&= P\left(\frac{-0.002}{\sigma/\sqrt{n}} \leq Z \leq \frac{0.002}{\sigma/\sqrt{n}}\right) \\
&= \Phi\left(\frac{0.002}{\sigma/\sqrt{n}}\right) - \Phi\left(-\frac{0.002}{\sigma/\sqrt{n}}\right) \quad (Z \text{ on jatk. sm.}) \\
&= 2\Phi\left(\frac{0.002}{\sigma/\sqrt{n}}\right) - 1 \quad (\text{symmetriaominaisuus}) \\
&\geq 0.9,
\end{aligned}$$

josta saadaan ehto

$$\Phi\left(\frac{0.002}{\sigma/\sqrt{n}}\right) \geq 0.95.$$

Normaalijakauman taulukosta saadaan

$$\frac{0.002}{\sigma/\sqrt{n}} \geq 1.645 \Rightarrow \sqrt{n} \geq \frac{1.645 \cdot 0.004}{0.002} \Rightarrow n \geq 11,$$

eli on otettava vähintään $n = 11$ näytettä halutun tarkkuuden saavuttamiseksi.

6. Lentoyhtiö tietää kokemuksesta, että keskimäärin 5 % paikan varanneista jää saapumatta koneeseen. Niinpä yhtiö myykin 184 lippua koneeseen, johon mahtuu 180 matkustajaa. Oletetaan, että paikan varaajat ovat toisistaan riippumattomia. Laske todennäköisyys, että jokainen lennolle todella saapuva saa paikan

a) tarkasti.

b) Poisson-jakauman avulla.

c) käyttämällä normaalijakauma-approksimaatiota jatkuvuuskorjauksella ja ilman.

Ratkaisu:
Tarkastellaan lennolle tulevien sijaan lennolta poisjäävien matkustajien lukumäärää. Olkoon siis

$$X = \text{"lennolta poisjäävien lukumäärä"}.$$

a) Nyt siis tarkastellaan jonkin "suotuisan" tapahtuman esiintymiskertojen lukumäärää toistokokeessa (jos lennolta poisjääminen katsotaan suituisaksi), joten $X \sim \mathrm{Bin}(184, 0.05)$. Kaikki saa paikan, jos poisjääviä on vähintään 4, joten kysytään todennäköisyyttä

$$P(X \geq 4).$$

Riittävän edistyneelle laskimelle todennäköisyyden laskeminen suoraan ei tuota mitään ongelmia. Jos kuitenkin käytetään funktiolaskinta, kannattaa hyödyntää komplementtitapahtumaa seuraavasti:

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) = 1 - P(X \leq 3) \quad \text{(sillä } X \text{ saa vain kokonaislukuarvoja)} \\ &= 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) \approx 0.984. \end{aligned}$$

b) Koska $np = 184 \cdot 0.05 = 9.2$, niin $X \sim \mathrm{Poi}(9.2)$ likimain. Jälleen komplementtitapahtuman avulla todennäköisyys voidaan laskea käsin ja saadaan

$$P(X \geq 4) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) \approx 0.982.$$

c) Ilman jatkuvuuskorjausta saadaan

$$\begin{aligned} P(X \geq 4) = 1 - P(X < 4) &\overset{\text{stand.}}{=} 1 - P\left(\frac{X - np}{\sqrt{np(1-p)}} < \frac{4 - 9.2}{\sqrt{9.2 \cdot 0.95}}\right) \\ &= 1 - \Phi\left(-\frac{5.2}{\sqrt{9.2 \cdot 0.95}}\right) \\ &= \Phi\left(\frac{5.2}{\sqrt{9.2 \cdot 0.95}}\right) \quad \text{(symmetriaominaisuuden nojalla)} \\ &\approx 0.961. \quad \text{(taulukosta)} \end{aligned}$$

Jatkuvuuskorjauksella saadaan

$$P(X \geq 4) = P(X \geq 3.5) = \cdots = \Phi\left(\frac{5.7}{\sqrt{9.2 \cdot 0.95}}\right) \approx 0.973. \quad \text{(taulukosta)}$$

7. Rahoituskriisiin ajautunut professori Tuhattaito päätti sijoittaa jäljelle jäävän omaisuutensa arvopapereihin toivoen pikaista kurssin nousua ja siten äkillistä rikastumista. Tietyn arvopaperin hinta n viikon kuluttua on satunnaismuuttuja $S(n)$. Suositun arvopaperimallin mukaan peräkkäisten viikkojen hintojen suhde $\frac{S(n)}{S(n-1)}$ noudattaa log-normaalijakaumaa eli $\ln(S(n)/S(n-1)) \sim N(\mu, \sigma^2)$. Edelleen malli olettaa, että peräkkäisten viikkojen hintasuhteet ovat riippumattomia. Tilastollisen aikasarja-analyysin perusteella Tuhattaito oli onnistunut määrittämään jakauman parametrien arvoiksi $\mu = 0.0165$ ja $\sigma = 0.0730$.

a) Millä todennäköisyydellä ostohetkestä $t = 0$ seuraavan kahden viikon aikana arvopaperin hinta nousee yli alkupääoman $S(0)$?

b) Mikä on sijoitetun omaisuuden arvon odotusarvo vuoden kuluttua?

c) Laske normaalijakauma-approksimaatiolla todennäköisyys, että hinta on vuoden kuluttua välillä $[2S(0), 4S(0)]$. Mikä on tarkka todennäköisyys?

Vihje: Seuraavat tulokset voivat olla avuksi.

(i) $\ln(ab) = \ln a + \ln b$,

(ii) Jos X, Y ovat riippumattomia sm:ia, niin $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$.

(iii) $E(X) = \exp(\mu + \frac{\sigma^2}{2})$, kun $\ln X \sim N(\mu, \sigma^2)$.

(iv) Jos $X \sim N(\mu_1, \sigma_1^2)$ ja $Y \sim N(\mu_2, \sigma_2^2)$, niin $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Ratkaisu: Merkitään $X_n = \ln(S(n)/S(n-1)) \sim N(\mu, \sigma^2)$.

a) Kysytään todennäköisyyttä $P(S(2) > S(0))$. Koska hintojen suhteen jakautuminen tiedetään, pyritään aikaansaamaan suhteita. Arvopaperin hinta 2 viikon kuluttua voidaan kirjoittaa muodossa

$$S(2) = \frac{S(2)}{S(1)}\frac{S(1)}{S(0)}S(0),$$

joten

$$S(2) > S(0) \Leftrightarrow \frac{S(2)}{S(1)}\frac{S(1)}{S(0)} > 1 \overset{\text{puolittain log.}}{\Rightarrow} \ln\left(\frac{S(2)}{S(1)} \cdot \frac{S(1)}{S(0)}\right) > 0,$$

sillä $\ln 1 = 0$ ja logaritmi on aidosti kasvava funktio. Vihjeen (i) mukaan kysytään todennäköisyyttä

$$P(S(2) > S(0)) = P(X_1 + X_2 > 0).$$

Vihjeen (iv) mukaan $X_1 + X_2 \sim N(2\mu, 2\sigma^2)$. Standardisoimalla $X_1 + X_2$ saadaan

$$\begin{aligned} P(S(2) > S(0)) =& P\left(\frac{X_1 + X_2 - \mathrm{E}(X_1 + X_2)}{\sigma_{X_1+X_2}} > \frac{0 - 2\mu}{\sqrt{2}\sigma}\right) \\ \overset{\text{symm.}}{=}& \Phi\left(\sqrt{2}\frac{\mu}{\sigma}\right) \approx \Phi(0.32) \\ \overset{\text{taul.}}{\approx}& 63\%. \end{aligned}$$

b) Käytetään samaa ideaa kuin a)-kohdassa ja esitetään $S(52)$ osamääriena avulla seuraavasti

$$S(52) = \frac{S(52)}{S(51)}S(51) = \cdots = Y(52)Y(51)\dots Y(1)S(0),$$

missä on merkitty $Y(n) = \frac{S(n)}{S(n-1)}$. Koska $Y(n)$:t ovat samalla tavalla jakautuneita ja riippumattomia muuttujia sekä $\ln Y(n) \sim N(\mu, \sigma^2)$, niin vihjeen (iii) mukaan

$$\mathrm{E}(S(52)) = \mathrm{E}(Y(1))^{52}S(0) = \exp(0.0191645)^{52}S(0) \approx 2.7S(0)$$

c) Käytetään samaa ideaa kuin b)-kohdassa, jolloin

$$2S(0) \leq S(52) \leq 4S(0) \Leftrightarrow 2 \leq Y(1)Y(2)\ldots Y(52) \leq 4.$$

Ottamalla puolittain logaritmi saadaan vihjeen (i) mukaan

$$\ln 2 \leq \sum_{n=1}^{52} X_n \leq \ln 4.$$

Koska X_n:t ovat samalla tavalla jakautuneita ja $\mathrm{E}(X_n) = \mu$, niin summan odotusarvo on 52μ. Vastaavalla tavalla saadaan summan hajonnaksi $\sqrt{52}\sigma$. Standardisoidaan summa, jolloin normaalijakauma-approksimaation mukaan

$$Z = \frac{\sum_{n=1}^{52} X_n - 52\mu}{\sqrt{52}\sigma} \overset{\text{likim.}}{\sim} N(0,1).$$

Kysytään todennäköisyyttä

$$\begin{aligned} P\left(\frac{\ln 2 - 52\mu}{\sqrt{52}\sigma} \leq Z \leq \frac{\ln 4 - 52\mu}{\sqrt{52}\sigma}\right) &\approx \Phi(1.00) - \Phi(-0.31) \\ &= \Phi(1.00) + \Phi(0.31) - 1 \quad \text{(symmetria-ominaisuus)} \\ &\approx 46\%. \quad \text{(taulukosta)} \end{aligned}$$

Edellä muuttujan Z jakaantumislaki on itse asiassa tarkka vihjeen (iv) mukaan, joten normaalijakauma-approksimaatiota ei olisi tarvinnut käyttää lainkaan ja saatu todennäköisyys on tarkka.

8. Mahtavat Masiinat Oy on kehittänyt elektronisia laitteita, jotka toimivat toisistaan riippumatta 95% todennäköisyydellä. Laitteet laivataan 400 kappaleen laatikoissa.

   a) Millä todennäköisyydellä laatikossa on vähintään 390 toimivaa laitetta?

   b) Yhtiö antaa takuun, että kussakin laatikossa vähintään k laitetta toimii. Mikä on suurin k, jolla takuu täyttyy vähintään 95% todennäköisyydellä? Käytä normaalijakauma-approksimaatiota jatkuvuuskorjauksella.

Ratkaisu:
Laite on toimiva todennäköisyydellä $p = 0,95$. Merkitään satunnaismuuttujalla X laatikossa olevien toimivien laitteiden lukumäärää. $X \sim \mathrm{Bin}(n, p)$, missä $n = 400$.

a) Voidaan laskea tarkasti

$$P(X \geq 390) = \sum_{390}^{400} \binom{400}{k} p^k (1-p)^{400-k} \simeq \underline{0,00940}$$

b) Approksimoiden on $X \sim \mathrm{N}(np, np(1-p)) = \mathrm{N}(380, 19)$, jolloin

$$\begin{aligned} q_k &= P(X \geq k) = 1 - P(X < k) \\ &= 1 - \Phi(\frac{k-380}{\sqrt{19}}) = \Phi(\frac{380-k}{\sqrt{19}}) \end{aligned}$$

Nyt vaaditaan, että $q_k \geq 0,95$:

$$\begin{aligned} \Phi(\frac{380-k}{\sqrt{19}}) &\geq 0,95 \\ \frac{380-k}{\sqrt{19}} &\geq 1,645 \quad \text{Normaalijakauman taulukosta} \\ k &\leq 372,8 \end{aligned}$$

Jatkuvuuskorjauksella tämä on $k - \frac{1}{2} \leq 372,8$, joten saadaan suurin $\underline{k = 373}$.

9. Fyysikot käyttävät satunnaiskävelyä diffuusion tai yleisesti partikkelien satunnaisliikkeen mallintamiseen. Partikkeli suorittaa siirtymiä ajanhetkillä $t = 1, 2, \ldots$ ja sen paikka S_n ajanhetkellä n voidaan ajatella olevan siirtymien $X_1, \ldots, X_n$ summa. Oletetaan, että siirtymät ovat riippumattomia ja samalla tavalla jakautuneita niin, että partikkeli voi samalla todennäköisyydellä joko pysyä paikallaan tai siirtyä yhden yksikön verran $(+1)$ oikealle tai vasemmalle (-1). Määrää normaalijakauma-approksimaatiolla todennäköisyys, että 10000 siirtymän jälkeen partikkeli on vähintään 100 yksikköä lähtöpisteestä oikealle.

Ratkaisu:
Olkoon X_i partikkelin siirtymä hetkellä i.
Tällöin $X_i \in \{-1, 0, 1\}$ ja $P(X_i = -1) = P(X_i = 0) = P(X_i = 1) = \frac{1}{3}$.

$$\begin{aligned}
E(X_i) &= (-1) \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0 \\
E(X_i^2) &= (-1)^2 \cdot \frac{1}{3} + 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{1}{3} = \frac{2}{3} \\
\mathrm{Var}(X_i) &= E(X_i^2) - E(X_i)^2 = \frac{2}{3} - 0^2 = \frac{2}{3}
\end{aligned}$$

Kappaleen paikka hetkellä k on $S_k = \sum_{i=1}^{k} X_i$ ja voidaan laskea sille

$$\begin{aligned}
E(S_k) &= E(\sum_{i=1}^{k} X_i) = \sum_{i=1}^{k} E(X_i) = 0 \\
\mathrm{Var}(S_k) &= \mathrm{Var}(\sum_{i=1}^{k} X_i) = \sum_{i=1}^{k} \mathrm{Var}(X_i) = k \cdot \frac{2}{3}
\end{aligned}$$

Hetkellä $k = 10000$ on $S_{10000} = S \overset{\text{likim.}}{\sim} \mathrm{N}(0, 10000 \cdot \frac{2}{3})$ ja todennäköisyys, että kappale on vähintään 100 yksikköä lähtöpisteestä oikealla käyttäen jatkuvuuskorjausta on

$$\begin{aligned}
P(S \geq 100 - \frac{1}{2}) &= 1 - P(S < 99,5) \\
&= 1 - P\left(\frac{S - 0}{100\sqrt{\frac{2}{3}}} < \frac{99,5 - 0}{100\sqrt{\frac{2}{3}}}\right) \\
&= 1 - \Phi\left(0,995 \cdot \sqrt{\frac{3}{2}}\right) \\
&= 1 - \Phi(1,21862) \simeq 1 - 0,888506 \\
&= \underline{0,111}
\end{aligned}$$