



Stockholms
universitet

Inspektionsparadoxen och observationsplaner

Emmi Sarinko

Kandidatuppsats 2010:5
Matematisk statistik
Juni 2010

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Inspektionsparadoxen och observationsplaner

Emmi Sarinko*

Juni 2010

Sammanfattning

Vissa urvalsmetoder, eller observationsplaner, resulterar i att urval görs med sannolikheter proportionella mot något relevant mått. Detta kan leda fram till situationer som känns paradoxala och även leda till felaktiga analyser av något givet datamaterial. Det paradoxala bakomliggande fenomenet kallas inspektionsparadoxen. Inspektionsparadoxen uppstår inom förnyelseteori och köteori och har även på naturligt sätt en mening i diskreta urvalssituationer. Vi diskuterar generella mekanismer som ger upphov till inspektionsparadoxen och presenterar även ett antal illustrerande simuleringar och beräkningar.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige. E-post: emmisarinko@hotmail.com. Handledare: Åke Svensson.

Inspektionsparadoxen och observationsplaner

Emmi Sarinko

Sammanfattning

Vissa urvalsmetoder, eller observationsplaner, resulterar i att urval görs med sannolikheter proportionella mot något relevant mått. Detta kan leda fram till situationer som känns paradoxala och även leda till felaktiga analyser av något givet datamaterial. Det paradoxala bakomliggande fenomenet kallas inspektionsparadoxen.

Inspektionsparadoxen uppstår inom förnyelseteori och köteori och har även på naturligt sätt en mening i diskreta urvalssituationer.

Vi diskuterar generella mekanismer som ger upphov till inspektionsparadoxen och presenterar ett antal illustrerande simuleringar och beräkningar.

Abstract

Certain methods of selection, or observation plans, is such that selection is made with probabilities proportional to some given measure. This can lead to situations that feel paradoxal and one can easily make mistakes in a statistical analysis of a given data material. The paradoxal phenomenon behind everything is called the inspection paradox.

The inspection paradox arise within renewal theory and queueing theory and has also in a natural way a meaning in discrete sampling situations.

We discuss general mechanisms behind the inspection paradox and present a number of descriptive simulations and computations.

Förord

Denna uppsats utgör ett examensarbete om 15 högskolepoäng och leder till en kandidatexamen i matematisk statistik vid Matematiska institutionen, Stockholms universitet.

Jag vill tacka professor Åke Svensson, min handledare på Matematiska institutionen, Stockholms universitet, för vägledning och goda råd under arbetets gång.

Innehåll

1	Introduktion	1
1.1	Inledning och bakgrund	1
1.2	Syfte och metod	1
2	Mekanismer bakom inspektionsparadoxen	2
2.1	Diskreta fall	2
2.2	Förnyelseprocesser	5
2.2.1	Ålder och återstående tid	7
2.2.2	Förnyelse-belöningsprocess	7
2.3	M/G/1-kö	9
3	Tillämpningar och observationsplaner	10
3.1	Bollar i lådor	10
3.2	Liftarens väntetid	12
3.3	Hastighetsmätningar	13
3.4	Sjukdomsspårning	14
3.5	Köer	18
3.5.1	Alltid sysselsätt system	18
3.5.2	Väntetid i M/G/1-kö	19
4	Diskussion	25

1 Introduktion

1.1 Inledning och bakgrund

Antag att bussarna avgår från busshållplatsen, där du bor, med fem-minuters intervall i rusningstrafik. Varje morgon vid 07:00 ställer du dig vid hållplatsen och väntar. Varför känns det som om du alltid får vänta på bussen längre än 2.5 minuter? Förklaringen finns inom förnyelse teorin. Man kan visa att fördelningen för längden av ett förnyelseintervall, till exempel tiden mellan två bussar, som innefattar en viss tidpunkt t skiljer sig från fördelningen för längden av ett slumpvist valt förnyelseintervall. Väntevärdet för längden av det intervall som innehåller punkten t är stokastiskt längre än väntevärdet för längden av övriga intervall.

Vissa urvalsmetoder, eller observationsplaner, gör att urval görs med sannolikheter proportionella mot något relevant mått. *Length-biased sampling* bygger på principen att sannolikheter är proportionella mot längder av vissa slumpmässigt valda intervall [10]. Om vi till exempel väljer en viss punkt på tidsaxeln och tänker oss att tidsaxeln är övertäckt av disjunkta intervall, så är sannolikheten att punkten kommer från ett visst intervall proportionell mot intervallens längd. *Size-biased sampling* lyder analogt: stickprov är proportionella mot populationens storlek. Som exempel kan man tänka sig att vi har ett antal urnor innehållande kulor. Om vi blandar alla kulor i en påse och väljer en kula slumpvis, så är sannolikheten att kulan kommer från en viss urna proportionell mot det totala antalet kulor som kommer från den urnan.

Urvalsmetoder som de just beskrivna leder till oväntade situationer som kan kännas paradoxala. Därför har fenomenet fått namnet *inspektionsparadoxen*.

I avsnitt 2 börjar vi med att studera klasstorlekar och presenterar två metoder för att beräkna väntevärden i diskreta fall. Vi kommer konstatera att de två metoderna ger olika väntevärden. Vi diskuterar förnyelseprocesser och tillämpningar som vi kan utnyttja i samband med inspektionsparadoxen. I slutet av avsnittet studerar vi hur köteori relaterar till inspektionsparadoxen. I avsnitt 3 presenterar vi tillämpningar som är baserade på teorin från avsnitt 2. Tillämpningarna är hämtade mer eller mindre från vardagliga situationer. Vi illustrerar exemplen med beräkningar och genom simuleringar. Avsnitt 4 är en kort diskussion kring uppsatsen.

1.2 Syfte och metod

Syftet med denna uppsats är att beskriva generella mekanismer, vilka ger upphov till inspektionsparadoxen. Vi hoppas kunna ge läsaren en så bred bild som möjligt av fenomenet inspektionsparadoxen. För att komplettera och illustrera teorin presenterar vi ett antal beräkningar och simuleringar.

2 Mekanismer bakom inspektionsparadoxen

En gemensam nämnare för situationer, där inspektionsparadoxer uppstår, är att olika observationsplaner ger olika fördelningar för observationerna. Vi börjar med att studera detta i det diskreta fallet.

2.1 Diskreta fall

1982 publicerades en artikel, "Why Your Classes Are Larger Than Average", [4], skriven av David Hemenway. Författaren studerar fixa klasstorlekar i skolan och jämför medelvärdet av klasstorlek, \bar{X} , med väntevärdet av klasstorlek, μ_X .

Vi kan beskriva Hemenways resultat enligt följande: Låt M , $M \geq 1$, vara ett antal personer som delas upp i n stycken olika stora klasser. Låt X_i , $i \in \{1, \dots, n\}$, beteckna storleken av den i :te klassen. Då har vi att $M = \sum_{i=1}^n X_i$ och väntevärdet för storleken av en slumpmässigt vald klass ("medelvärdet av klasstorlek") är

$$\bar{X} = \frac{M}{n}.$$

Vi antar att alla elever har samma sannolikhet att bli utvald, då vi slumpmässigt väljer en elev i skolan. Enligt välkända räkningar får vi väntevärdet för storleken av den klass som innehåller den utvalda eleven

$$\mu_X = \sum_{i=1}^n \frac{X_i}{M} \cdot X_i = \sum_{i=1}^n \frac{X_i^2}{M}.$$

Skillnaden mellan medelvärdet och väntevärdet av klasstorleken blir således

$$\begin{aligned} \mu_X - \bar{X} &= \frac{n \cdot \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}{(M \cdot n)} = \\ &= \frac{n \cdot \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}{n^2} \cdot \frac{n}{M} = \frac{\sigma_X^2}{\bar{X}}, \end{aligned} \quad (1)$$

där σ_X^2 är klasstorlekens varians.

Vi observerar att skillnaden mellan väntevärdet, μ_X , och medelvärdet, \bar{X} , är proportionell mot klasstorlekens varians och medelvärde. Vi ser vidare att $\mu_X \geq \bar{X}$, och likhet gäller endast när alla klasser är lika stora.

Stein och Dattero, [3], refererar till Hemenways skolklass-exempel och studerar fallet då klasstorlekarna är stokastiska. Stein och Dattero vill jämföra väntevärden för klassernas storlek då man slumpmässigt väljer en klass i skolan jämfört med om man slumpmässigt väljer en elev i skolan och betraktar storleken på klassen som denne elev tillhör. Stein och Dattero definierar på så sätt två urvalsmetoder och visar att väntevärdena blir olika beroende på vilken metod man väljer. Vi tänker oss klasserna som lådor. Lådorna skulle lika bra kunna innehålla ett slumpmässigt antal partiklar, bollar, kakor, eller vad vi nu är intresserade av. Vi antar att antalet saker i de respektive lådorna ges av oberoende och lika fördelade slumpvariabler X_i och antar att alla slumpvariabler

har samma kända sannolikhetsfunktion $p_j = P(X_i = j)$, $j \geq 1$. Vi kommer låta antalet lådor n , $n \geq 1$, gå mot oändligheten. Vi accepterar inte tomma lådor, så det följer att det totala antalet saker, M , går mot oändligheten då antalet lådor går mot oändligheten.

Metod 1: Vi väljer slumpmässigt en låda, låda i säg. Väntevärdet för antalet föremål i den i :te lådan betecknas med μ_X och variansen med σ_X^2 . Därav följer att

$$E[X_i] = \sum_{j=1}^{\infty} j p_j = \mu_X \quad \text{och} \quad \sigma_X^2 = \sum_{j=1}^{\infty} j^2 p_j - \mu_X^2. \quad (2)$$

Metod 2: Vi blandar först ihop samtliga föremål och väljer sedan slumpmässigt ett föremål. Låt X_I vara antalet föremål i lådan det valda föremålet kommer från, det vill säga X_I är antal saker i lådan inklusive den valda saken. Observera att detta indirekt är ett sätt att välja en låda I . Givet antal saker i en låda blir lådan på så sätt vald proportionellt mot innehållet. Då vi låter antal lådor n gå mot oändligheten får vi enligt Stein och Dattero att

$$\lim_{n \rightarrow \infty} P(X_I = j) = \frac{j p_j}{\mu_X}, \quad j \geq 1, \quad (3)$$

och

$$E[X_I] = \sum_{j=1}^{\infty} j P(X_I = j) = \sum_{j=1}^{\infty} \frac{j^2 p_j}{\mu_X} = \frac{\sigma_X^2 + \mu_X^2}{\mu_X} = \mu_X + \frac{\sigma_X^2}{\mu_X}. \quad (4)$$

Skillnaden mellan väntevärdet för låda i och I är

$$E[X_I] - E[X_i] = \frac{\sigma_X^2}{\mu_X}, \quad (5)$$

vilket liknar Hemenways resultat (1). Med hjälp av (4) får vi att

$$\begin{aligned} E\left[\frac{1}{X_I}\right] &= \sum_{j=1}^{\infty} \frac{1}{j} P(X_I = j) = \sum_{j=1}^{\infty} \frac{1}{j} \frac{j p_j}{\mu_X} = \frac{1}{\mu_X} \\ &\iff \\ &\frac{1}{E\left[\frac{1}{X_I}\right]} = \mu_X. \end{aligned} \quad (6)$$

Vi ser att vänsterledet i (6) är ett harmoniskt medelvärde av X_I . Antag att vi har ett stickprov med observationer x_i , $i \in \{1, \dots, n\}$, som har samlats enligt metod 2. Harmoniska medelvärdet av stickprovet är

$$HM = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}.$$

Följande definition härstammar från Hus och Zhuangs artikel [9].

Definition 2.1. Låt X_1 och X_I vara två slumpvariabler. Vi säger att X_1 är mindre än X_I

- i stokastisk ordning om $E[X_1] \leq E[X_I]$;
- i likelihoodkvotordning om för X_1 :s och X_I :s sannolikhetsfunktioner, $P(X_1 = j)$ respektive $P(X_I = j)$, gäller att $\frac{P(X_I=j)}{P(X_1=j)}$ är växande i j , där kvoten är definierad.

Ross [6] sammanfattar i en sats inte bara teorin som Stein och Dattero använder utan även flera andra samband som är användbara, då man diskuterar inspektionsparadoxen.

Sats 2.2. Med samma beteckningar som ovan gäller att

- (i) X_I är större än X_1 i likelihoodkvotordning.
- (ii) $P(X_I \geq r) \geq \sum_{j \geq r}^{\infty} p_j$ i stokastisk ordning.
- (iii) $P(X_I = j) \geq \frac{jp_j}{\mu_X} \cdot \frac{n\mu_X}{j+(n-1)\mu_X}$.
- (iv) $\lim_{n \rightarrow \infty} P(X_I = j) = \frac{jp_j}{\mu_X}$.

Bevis. Genom att använda Bayes sats får vi att

$$P(X_I = j | I = k) = P(X_k = j | I = k) = \frac{P(I = k | X_k = j)p_j}{P(I = k)}. \quad (7)$$

Från täljaren har vi att

$$P(I = k | X_k = j, X_i, i \neq k) = \frac{j}{j + \sum_{i \neq k} X_i},$$

vilket innebär att

$$P(I = k | X_k = j) = E\left[\frac{j}{j + \sum_{i \neq k} X_i} \mid X_k = j\right] = E\left[\frac{j}{j + \sum_{i \neq k} X_i}\right]. \quad (8)$$

Ekvation (8) och det faktum att $P(I = k) = \frac{1}{n}$ ger att (7) kan skrivas

$$P(X_I = j | I = k) = jp_j E\left[\frac{n}{j + \sum_{i \neq k} X_i}\right].$$

Då slumpvariablerna är lika fördelade beror högerledet i ovanstående ekvation inte på k , vilket ger

$$P(X_I = j) = jp_j E\left[\frac{n}{j + \sum_{i \neq k} X_i}\right], \quad j \geq 1, \quad (9)$$

\iff

$$\frac{P(X_I = j)}{p_j} = nE\left[\frac{1}{1 + \sum_{i \neq k} \frac{X_i}{j}}\right]. \quad (10)$$

Högerledet i (10) är en växande funktion i j , så del (i) stämmer. Del (ii) stämmer genom (i), eftersom likelihoodkvotordning är starkare än stokastisk ordning, se [9]. Tillämpning av Jensens olikhet¹ ger att högerledet i (9) blir

$$P(X_I = j) \geq jp_j \frac{n}{j + (n-1)\mu_X},$$

vilket är del (iii). Del (iv) fås dels då vi tillämpar stora talens lag i (9),

$$\frac{1}{\frac{j}{n} + \sum_{i \neq k} \frac{X_i}{n}} \rightarrow \frac{1}{\mu_X}, \quad \text{då } n \rightarrow \infty,$$

och dels genom Lebesgues dominerade konvergenssats (se till exempel [8]). \square

Angus [1] använder Ross's sats och utvecklar den litet: Av ekvation (9) får vi

$$E[X_I] = E\left[\frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i}\right], \quad (11)$$

vilket är ett enkelt sätt att skatta väntevärdet för storleken av lådan då vi väljer ett föremål, då vi vet hur många föremål varje låda innehåller. Beviset för detta är en liten räkning:

$$\begin{aligned} E[X_I] &= \sum_{j \geq 1} j^2 p_j E\left[\frac{n}{j + \sum_{i \neq k} X_i}\right] = \sum_{j \geq 1} j^2 p_j E\left[\frac{n}{j + \sum_{i=2}^n X_i}\right] = \\ &= \sum_{j \geq 1} j^2 p_j E\left[\frac{n}{\sum_{i=1}^n X_i} \mid X_I = j\right] = \sum_{j \geq 1} p_j E\left[\frac{n X_I^2}{\sum_{i=1}^n X_i} \mid X_I = j\right] = \\ &= nE\left[\frac{X_I^2}{\sum_{i=1}^n X_i}\right] = E\left[\frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i}\right]. \end{aligned}$$

2.2 Förnyelseprocesser

I många praktiska situationer är teorin för förnyelseprocesser grunden för statistiska modeller. Vi kan översätta Steins och Datteros låd-metoder till tidsintervall och förnyelse teori och vice versa. Vi tänker oss att lådor är tidsintervall och att enskilda saker är tidpunkter i tidsintervallen.

Definition 2.3. En stokastisk process $\{N(t), t \geq 0\}$, sägs vara en räkneprocess om $N(t)$ representerar "antalet händelser" som inträffat till och med tidpunkten t .

¹Låt $(\Omega, \mathcal{F}, \mathbb{P})$ vara ett sannolikhetsrum och X en reellvärd stokastisk variabel på Ω . Om väntevärdet $E[X] = \int_{\omega \in \Omega} |X|(\omega) d\mathbb{P}(\omega)$ är ändligt och $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ är en konvex funktion så gäller olikheten $\varphi(E[X]) \leq E[\varphi(X)]$.

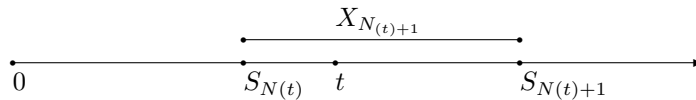
Allmänt gäller för en räkneprocess $N(t)$ att (se Ross, [7]):

- (i) $N(t) \geq 0$.
- (ii) $N(t)$ är ett heltal för alla t .
- (iii) Om $s \leq t$, så $N(s) \leq N(t)$.
- (iv) För $s \leq t$, så är $N(t) - N(s)$ antalet händelser i intervallet $(s, t]$.

Förnyelseprocesser är en speciell sorts räkneprocesser:

Definition 2.4. Låt $\{N(t), t \geq 0\}$ vara en räkneprocess och antag att tiden, X_i , mellan den $i - 1$:te och den i :te händelse ges av en följd av icke-negativa stokastiska variabler $\{X_i\}_{i \geq 1}$ som är oberoende och har samma fördelning. Då sägs räkneprocessen $\{N(t), t \geq 0\}$ vara en förnyelseprocess.

Vi betecknar med $S_{N(t)}$ tidpunkten för händelsen som observerats senast innan eller vid tidpunkten t . Följande bild illustrerar våra beteckningar.



Figur 1: Tidsintervall

Vi låter väntevärde och varians för X_i vara $E[X_i] = \mu_X$ respektive $\text{var}[X_i] = \sigma_X^2$. Man kan nu visa inspektionsparadoxen. För något fixt t betraktar vi tidsintervallet

$$X_{N(t)+1} = S_{N(t)+1} - S_{N(t)}.$$

Genom att betinga fördelningen för $X_{N(t)+1}$ på föregående tidsintervall vid tiden t har vi att, se [7],

$$P(X_{N(t)+1} > x | S_{N(t)} = t - s) = E[P(X_{N(t)+1} > x | S_{N(t)} = t - s)].$$

Vi får

$$P(X_{N(t)+1} > x | S_{N(t)} = t - s) = \begin{cases} 1 & \text{om } x < s \\ \frac{1-F(x)}{1-F(s)} & \text{om } x \geq s \end{cases}.$$

Alltså har vi att $P(X_{N(t)+1} > x | S_{N(t)} = t - s) \geq 1 - F(x)$ för alla s . Genom att ta väntevärdet av båda leden i olikheten får vi

$$P(X_{N(t)+1} > x) \geq P(X_1 > x) = 1 - F(x) \quad \text{för alla } x \geq 0,$$

vilket medför att $E[X_{N(t)+1}] \geq E[X_1]$ (se [5], [7] och jämför med Sats 2.2 (ii)). Detta visar att fördelningen för $X_{N(t)+1}$ är stokastisk större än fördelning för X_1 . Det vill säga, ett förnyelseintervall som innehåller en fix punkt t är stokastiskt längre än ett godtyckligt förnyelseintervall.

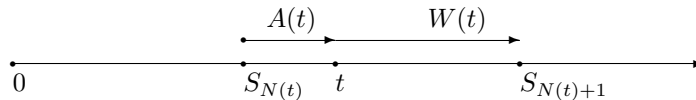
2.2.1 Ålder och återstående tid

Slutsatsen från föregående sektion kan även nås genom att använda ålder och återstående tid för att analysera förnyelseintervallen i en förnyelseprocess. Låt $N(t)$ vara en förnyelseprocess och beteckna som tidigare med $S_{N(t)}$ tidpunkten för händelsen som vi har observerat senast innan eller vid tidpunkten t . Vi sätter

$$A(t) = t - S_{N(t)} \quad \text{och} \quad W(t) = S_{N(t)+1} - t.$$

$A(t)$ och $W(t)$ kallas *ålder* från senast händelse respektive *återstående tid* till nästa händelse. Notera att

$$X_{N(t)+1} = A(t) + W(t). \quad (12)$$



Figur 2: Ålder och återstående tid

Enligt Ross [7] är genomsnittlig ålder från senaste förnyelse och återstående tid till nästa förnyelse båda $\frac{E[X^2]}{2E[X]}$, då t blir godtyckligt stort. Av (12) får vi att väntevärdet för längden av förnyelseintervallen blir

$$\lim_{t \rightarrow \infty} E[X_{N(t)+1}] = \frac{E[X^2]}{2E[X]} + \frac{E[X^2]}{2E[X]} = \frac{E[X^2]}{E[X]} = \mu_X + \frac{\sigma_X^2}{\mu_X}, \quad (13)$$

vilket visar sig vara större än μ_X och sammanfaller med ekvation (4).

2.2.2 Förnyelse-belöningsprocess

Vi kan även betrakta förnyelseintervall och inspektionsparadoxen genom att använda belöningsprocesser: då vi väljer en fix tidpunkt t på tidsaxeln så kan vi visa att väntevärdet för belöningen under intervallet, som innehåller t , är större än belöningen under ett stokastiskt valt intervall.

Vi låter $\{N(t), t \geq 0\}$ vara en förnyelseprocess. Om vi observerar ett visst tidsintervall (eller en viss låda) så kan det vara av intresse att veta om intervallets längd (eller antal föremål i lådan) överstiger ett visst givet intressant mått. I det fallet längden (eller antalet) överstiger det givna måttet, så tänker vi oss att det sker en "belöning". Man kan till exempel tänka sig ett lotteri, där personen som "har intervall i " vinner ett pris om intervallet är tillräckligt långt.

Låt $\{X_n\}_{n \geq 1}$ vara en följd av positiva, oberoende och lika fördelade förnyelsetider med icke-aritmetisk sannolikhetsfunktion

$$F_X(x) = P(X_n \leq x), \quad n \geq 1.$$

Låt summan av intervalltiderna vara

$$S_n = \sum_{i=1}^n X_i, \quad S_0 = 0, \quad n \geq 1.$$

Vi betecknar med $Y_n, n \geq 1$, den belöning vi får under det n :te förnyelseintervallet och antar att paren $(X_n, Y_n), n \geq 1$, är oberoende och har samma fördelning med gemensam sannolikhetsfördelning

$$F_{XY}(x, y) = P(X_n \leq x \text{ och } Y_n \leq y), \quad n \geq 1.$$

Den marginella sannolikhetsfördelningen för Y_n är

$$F_Y(y) = P(Y_n \leq y), \quad n \geq 1.$$

Vi tillåter att Y_n kan bero på X_n (se [7]). Låt $Y(t)$ representera den totala mängden av belöningar till och med tiden t ,

$$Y(t) = \sum_{n=1}^{N(t)} Y_n, \quad t \geq 0, \quad Y_0 = 0.$$

Vidare låter vi väntevärden, kovariansen och varianser för X_n och Y_n vara

$$\begin{aligned} E[X_n] &= \mu_X, & E[Y_n] &= \mu_Y, \\ \text{cov}[X_n, Y_n] &= \sigma_{XY}, & \text{var}[X_n] &= \sigma_X^2, & \text{var}[Y_n] &= \sigma_Y^2. \end{aligned}$$

Följande resultat för förnyelse-belöningsprocesser är viktiga och bra att känna till, se [7] och [12].

Proposition 2.5. *Om $E[Y_n] < \infty$ och $E[X_n] < \infty$, då följer att*

$$(i) \lim_{t \rightarrow \infty} \frac{E[Y(t)]}{t} = \frac{\mu_Y}{\mu_X}.$$

$$(ii) \lim_{t \rightarrow \infty} \frac{E[Y_{N(t)+1}]}{t} = 0.$$

Då X_n och Y_n korrelerar med varandra, och vi låter t bli stort, så följer enligt Wilson [12] att

$$\lim_{t \rightarrow \infty} E[Y_{N(t)+1}] = \mu_Y + \frac{\sigma_{XY}}{\mu_X}. \quad (14)$$

Den intresserade läsaren kan se beviset till detta i Wilsons artikel. Återigen ser vi liknande resultat som tidigare. Ekvationen (14) skiljer sig från väntevärdet för belöningen om X_n och Y_n korrelerar med varandra.

2.3 M/G/1-kö

Köteori är intressant, eftersom den är baserad på enkla och realistiska situationer: kunder kommer oberoende av varandra till en betjäningstation, de väntar på sin tur i kön (då inga avhopp sker), kunderna blir betjänade under slumpmässigt lång tid och därefter de lämnar stationen. Det är lätt att tänka sig att då ankomstnivån ökar, bildas det en längre kö och väntetiden i kön blir längre. Det kan vara sant, men Cooper, Niu, och Srinivasan [2] menar att väntetiden i en kö kan bli godtyckligt lång även fast utnyttjandet av servicen är låg och väntevärdet för betjäningstiden förblir oförändrat. Författarna tar hjälp av *M/G/1*-köer och *Pollaczek-Khintchines formel* för att visa detta. En *M/G/1*-kö är en kö där, se [7],

- (i) Kunderna anländer enligt en poissonprocess med intensitet λ .
- (ii) Betjäningstiderna lyder generella fördelningen.
- (iii) Det finns en betjäningstation.

Betjäningssordningar kan se olika ut, men vi tänker oss här en betjäningssordning, där först inkommen blir först betjänad. Då gäller att kunden anländer antingen till tomma systemet och hans betjäning börjar direkt eller så anländer kunden då någon eller några finns i systemet och han får vänta. Väntetiden i kö, W_i , för kund nummer i är lika med arbete i systemet då kunden anländer. Pollaczek-Khintchines formel beskriver medelväntetiden i kö, W_Q , vilket även är genomsnittligt arbete i systemet. Betjäningstiden för kund nummer i betecknas med S_i . $E[S]$ och $\text{var}[S]$ är väntevärde respektive varians för betjäningstiden. Cooper, Niu och Srinivasan [2] visar att man kan skriva

$$W_Q = \frac{\rho}{1 - \rho} \frac{1}{2} \left(E[S] + \frac{\text{Var}[S]}{E[S]} \right), \quad (15)$$

där

$$\rho = \begin{cases} \lambda E[S] & \text{om } \lambda E[S] < 1 \\ 1 & \text{om } \lambda E[S] \geq 1 \end{cases}. \quad (16)$$

I ekvationen ovan står ρ för utnyttjandet av betjäningstation. Poängen med ekvation (15) är att man kan skriva W_Q som en produkt av medelkölängd och medeltid för återstående tid av den betjäning som pågår, då en kund kommer till systemet.

Då ρ är $\lambda E[S]$ får vi den bekanta formeln

$$W_Q = \frac{\lambda E[S^2]}{2(1 - \lambda E[S])},$$

vilken Ross använder i sin bok [7]. Ekvationen (15) visar tydligt att för alla positiva värden för ρ , oavsett hur små, kan W_Q bli godtyckligt stort. Detta är anmärkningsvärt och beror på length-biased sampling: termen $\frac{\text{var}[S]}{E[S]}$ kan bli stor, eftersom vi sannolikt kommer till systemet under en lång betjäning, se [2]. Denna kvot har vi diskuterat i tidigare sammanhang och det är vanligt att den uppstår i samband med inspektionsparadoxen.

3 Tillämpningar och observationsplaner

3.1 Bollar i lådor

Vi ger ett enkelt räkneexempel baserat på avsnitt 2.1. Exemplet visar hur Steins och Datteros låd-metoder fungerar, och vilken skillnad som kan uppstå beroende på vilken metod vi väljer att skatta antalet föremål i en låda.

Vi antar att vi har ett antal lådor som innehåller ett stokastiskt antal bollar. Vi låter X_i beteckna antalet bollar i respektive låda. Lådorna antas vara lika fördelade och oberoende av varandra.

Vi börjar med att betrakta fallet då vi har två lådor och då båda lådorna innehåller antingen en boll eller två bollar med lika stor sannolikhet, $P(X_i = j) = \frac{1}{2}$, där alltså $j \in \{1, 2\}$.

Enligt metod 1 väljer vi slumpmässigt en låda. Väntevärdet för antalet bollar i den lådan är

$$E[X_i] = \sum_{j=1}^2 jP(X_i = j) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2}.$$

Enligt metod 2 väljer vi slumpmässigt en boll. Låt X_I vara totala antalet bollar i den låda som innehåller den valda bollen. Vi får

$$\begin{aligned} P(X_I = 1) &= P(X_I = 1 | X_1 = 1, X_2 = 1) \cdot P(X_1 = 1, X_2 = 1) \\ &+ P(X_I = 1 | X_1 = 1, X_2 = 2) \cdot P(X_1 = 1, X_2 = 2) \\ &+ P(X_I = 1 | X_1 = 2, X_2 = 1) \cdot P(X_1 = 2, X_2 = 1) \\ &+ P(X_I = 1 | X_1 = 2, X_2 = 2) \cdot P(X_1 = 2, X_2 = 2) \\ &= 1 \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = \frac{5}{12}, \end{aligned}$$

och således även att $P(X_I = 2) = 1 - \frac{5}{12} = \frac{7}{12}$.

Väntevärdet för antal bollar i den lådan som innehåller den slumpmässigt valda bollen är

$$E[X_I] = \frac{5}{12} + 2 \cdot \frac{7}{12} = \frac{19}{12}.$$

Skillnaden mellan $E[X_I]$ och $E[X_i]$ är $\frac{1}{12}$.

Med samma antaganden som ovan ökar vi antal lådor till 20. Vi har fortfarande enligt metod 1 att $E[X_i] = 1.5$. Däremot har vi att $P(X_I = 1) \approx 0.341$ och $P(X_I = 2) \approx 0.659$, varav följer att väntevärdet enligt metod 2 blir $E[X_I] = 1.659$. I det här fallet blir skillnaden mellan $E[X_I]$ och $E[X_i]$ 0.159, vilket nästan är en fördubbling jämfört med fallet då vi hade 2 lådor.

$P(X_I = j)$	Antal lådor=2	Antal lådor=20
$P(X_I = 1)$	0.417	0.341
$P(X_I = 2)$	0.583	0.659
$E[X_i]$	1.5	1.5
$E[X_I]$	1.583	1.659
$E[X_I] - E[X_i]$	0.083	0.159

Tabell 1: Sannolikheter och väntevärden för antal bollar i lådan som innehåller den slumpmässigt valda bollen.

Det kan vara intressant att jämföra väntevärdena från de två olika metoderna, då vi har flera bollar. Först studerar vi fallet då antalet bollar är 10 och antalet lådor är 2. Väntevärdet för antal bollar i lådan blir enligt metod 1:

$$E[X_i] = \sum_{j=1}^{10} jP(X_i = j) = 1 \cdot \frac{1}{10} + \dots + 10 \cdot \frac{1}{10} = 5.5.$$

Med hjälp av sannolikheterna i tabell 2 nedan får vi att väntevärdet för antal bollar i lådan enligt metod 2 är

$$E[X_I] = \sum_{j=1}^{10} jP(X_I = j) = 6.3.$$

Vi ser att skillnaden mellan väntevärden enligt metod 1 och 2 blir 0.8.

I vårt sista fall låter vi antal bollar vara 10 men ökar antal lådor. Vi låter antal lådor bli tillräckligt stort för att kunna approximera sannolikheterna för varje fall enligt Sats 2.2 (iv). Då vi väljer en låda slumpmässigt får vi att väntevärdet för antal bollar i lådan är $E[X_i] = 5.5$. Då vi väljer en boll slumpmässigt får vi att väntevärdet för antal bollar i lådan är $E[X_I] = 7$.

Vi ser att skillnaden mellan dessa två väntevärden blir betydligt större då vi ökar antalet bollar och antalet lådor. Samtidigt ser vi en ökning i sannolikhet att vår slumpmässigt valda boll kommer från en låda med flera bollar. Detta är ett typiskt exempel på size-biased sampling, vilket leder till inspektionsparadoxen. Med en liten kontrollräkning kan vi även konstatera att delarna (ii) och (iii) i Sats 2.2 stämmer.

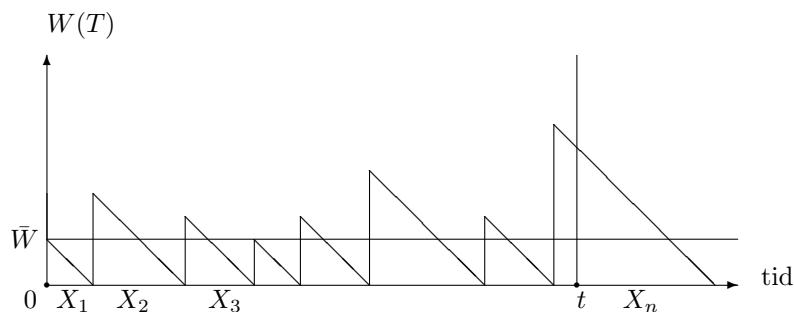
$P(X_I = j)$	Antal lådor=2	Antal lådor $\rightarrow \infty$
$P(X_I = 1)$	0.040	0.018
$P(X_I = 2)$	0.064	0.036
$P(X_I = 3)$	0.081	0.055
$P(X_I = 4)$	0.093	0.073
$P(X_I = 5)$	0.104	0.091
$P(X_I = 6)$	0.112	0.109
$P(X_I = 7)$	0.119	0.127
$P(X_I = 8)$	0.124	0.145
$P(X_I = 9)$	0.129	0.164
$P(X_I = 10)$	0.134	0.182
$E[X_i]$	5.5	5.5
$E[X_I]$	6.3	7.0
$E[X_I] - E[X_i]$	0.8	1.5

Tabell 2: Sannolikheter och väntevärden för antal bollar i lådan som innehåller den slumpmässigt valda bollen.

3.2 Liftarens väntetid

Vi tar ett exempel där vi använder oss av ålder och återstående tid. Vi antar oss veta att bilar kör på en landsväg enligt en poissonprocess med intensitet $\lambda = 10$ minuter. En man som bor vid vägen har observerat trafiken under en lång tid t . Vi har på så sätt tillgång till fixa tider för bilar som har passerat mannens hus. En liftare kommer och ställer sig vid vägkanten, där mannen bor, vid en slumpmässig tidpunkt $T \in (0, t]$. Vi antar att liftaren får skjuts av nästkommande bil. Hur lång tid behöver liftaren vänta tills det kommer en bil? Poissonprocessens ankomsttider är exponentialfördelade och typiskt för exponentialfördelningen är att händelserna är minneslösa. Av detta följer att kvarstående tid tills liftaren får skjuts är exponentialfördelad med intensitet $\frac{1}{10}$ och därmed blir förväntad väntetid för liftaren 10 minuter. Att tidsintervallet då liftaren anländer till vägkanten är längre än övriga tidsintervall är ett typiskt exempel på inspektionsparadoxen. Förklaringen är att det är större sannolikhet att liftaren anländer under ett längre tidsintervall än under ett kortare tidsintervall.

Väntetiden tills liftaren får skjuts, $W(T)$, är en funktion av tidpunkten T . Vi illustrerar situationen med ett sågkantsdiagram. Vi betecknar med X_i , $i \in \{1, \dots, n\}$, observerade tidsintervall mellan bilarna som passerar huset.



Figur 3: Liftarens väntetid

Vi känner redan till väntevärdet för väntetiden, men det kan vara intressant att även känna till liftarens genomsnittliga väntetid, \bar{W} . Enligt [11] är det sågkantsdiagrammets medelvärde och fås genom

$$\bar{W} = \frac{1}{t} \int_0^t W(T) dT \approx \frac{1}{t} \sum_{i=1}^n \frac{1}{2} X_i^2.$$

Då vi låter $t \rightarrow \infty$, har vi att genomsnittlig längd av ett tidsintervall blir

$$\bar{X} = \frac{t}{n} \quad \Leftrightarrow \quad t = \bar{X}n,$$

varav följer att

$$\bar{W} = \frac{1}{\bar{X}} \frac{1}{n} \sum_{i=1}^n \frac{1}{2} X_i^2 = \frac{1}{2} \frac{\bar{X}^2}{\bar{X}}.$$

För exponentialfördelningen gäller vid fixa tidsintervall att $E[X_i] = \frac{1}{\lambda} = \bar{X}$ och $\text{var}[X_i] = \frac{1}{\lambda^2} = \bar{X}^2$ [11]. Detta ger

$$\bar{X}^2 = (\bar{X})^2 + \text{var}[X_i] = (\bar{X})^2 + (\bar{X})^2 = 2(\bar{X})^2 \quad \implies \quad \bar{W} = \bar{X}.$$

Alltså, den genomsnittliga väntetiden för liftaren blir den genomsnittliga längden för intervallet. Detta medför att tidsintervallet, då liftaren anländer, blir längre än ett genomsnittligt intervall mellan två bilar.

3.3 Hastighetsmätningar

Trafiksimulering har varit av intresse under en lång tid, och ett intressant exempel på inspektionsparadoxen uppstår då vi undersöker hastigheter. Vi tänker oss att vi står vid vägkanten i punkten x och observerar hastigheter för bilar som passerar oss. Vägen är enkelriktad och bilarna startar samtidigt från slumpmässigt valda startpunkter och kör mot punkten x . Bilarna kör med konstanta hastigheter som är lika fördelade och oberoende av varandra.

Vi betecknar slumpmässiga bilhastigheter med U_i , $i \in \{1, \dots, n\}$. Hastigheterna är likformigt fördelade, antingen mellan 10 och 120 km/h ($U_i \sim U(10, 120)$) eller mellan 90 och 120 km/h ($U_i \sim U(90, 120)$). Vi simulerar hastigheter för bilar samt avståndet från de respektive bilarnas startpunkter till platsen vi står. Vi bestämmer en tid, t (timmar), som vi stannar vid vägkanten och mäter hastigheterna. Vi betecknar medelhastighet för bilar som kör förbi oss med \bar{U}_x och medelhastighet för alla bilar i vår experiment med \bar{U} . Då vi jämför \bar{U}_x och \bar{U} så märker vi att \bar{U}_x blir större än \bar{U} . Förklaringen till att inspektionsparadoxen uppstår här är, åtminstone intuitivt, enkel: Snabbare bilar blir observerade oftare än långsamma bilar, och under ett givet tidsintervall är sträckan proportionell mot hastigheten, se [3].

Vi illustrerar våra simuleringar i tabellerna nedan. Vi varierar hastigheter, antal bilar, och tiden som vi observerar bilar.

	(Antal bilar, observationstid)=(n, t)					
	(10, 2)	(100, 2)	(1000, 2)	(10, 10)	(100, 10)	(1000, 10)
\bar{U}_x	73.39	79.86	81.82	75.50	88.11	80.51
\bar{U}	66.63	65.70	65.46	67.24	65.78	65.31
$\bar{U}_x - \bar{U}$	6.76	14.16	16.36	8.26	22.33	15.20

Tabell 3: Medelhastigheter, $U_i \sim U(10, 120)$.

	(Antal bilar, observationstid)=(n, t)					
	(10, 2)	(100, 2)	(1000, 2)	(10, 10)	(100, 10)	(1000, 10)
\bar{U}_x	105.64	107.02	105.85	102.57	107.09	105.52
\bar{U}	104.70	105.65	105.17	101.67	106.20	104.76
$\bar{U}_x - \bar{U}$	0.94	1.37	0.68	0.90	0.89	0.76

Tabell 4: Medelhastigheter, $U_i \sim U(90, 120)$.

I tabell 3 ser vi att skillnaderna i medelhastighet är mindre då vi har få bilar (10 bilar) i vårt experiment. I tabell 4 är skillnaderna i medelhastighet betydligt större jämfört med i tabell 3, vilket naturligtvis beror på att variationen i hastigheterna är mindre i tabell 4. Skillnaden mellan de två medelhastigheterna blir mindre desto jämnare bilarna åker.

3.4 Sjukdomsspårning

I det här exemplet illustrerar vi en diskret förnyelse-belöningsprocess. Antag att vi vill spåra en smittsam sjukdom som plågar en skolas elever. Skolan är stor och barnen kommer från familjer av olika storlek. Vi tänker oss att om en familj har flera barn så går alla i samma skola. Då en del av skolans elever blivit sjuka, slår lärarna larm och beslutar spåra de familjer som har sjuka barn.

Ett slumpmässigt valt barn bär sjukdomen enligt bernoullifördelningen med sannolikhet p_s . Vi antar att barnen i skolan är lika fördelade, oberoende av

varandra och att alla barn har lika stor sannolikhet att vara smittade. Det betyder att om det finns ett sjukt barn i en familj så har syskonen inte större sannolikhet än andra barn att vara smittade. Låt $X_i, i \in \{1, \dots, n\}$, vara antal barn i skolan som kommer från familj i .

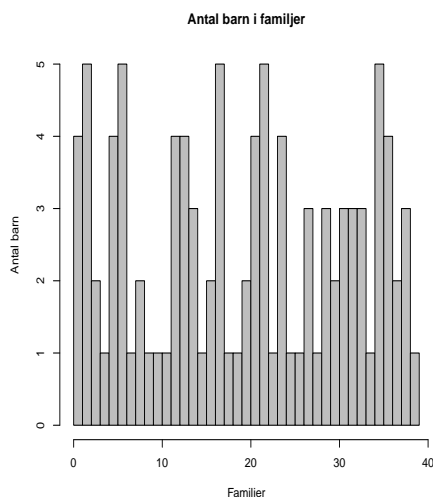
Vi numrerar samtliga barn i skolan och betecknar numret för ett barn vi väljer slumpmässigt med $j, j \geq 1$, och hans familj med X_n . Sannolikheten att välja ett barn j slumpmässigt är $p_j = \frac{1}{\sum_{i=1}^n X_i}$ och sannolikheten att välja en familj slumpmässigt är $\frac{1}{n}$.

Vi betraktar sjukdomen som en belöning: det föreligger belöning Y_n om smittan finns i familj X_n . Det är naturligt att tillåta att sannolikheten för Y_n beror av X_n (desto större familj, desto större sannolikhet att sjukdomen finns i familjen).

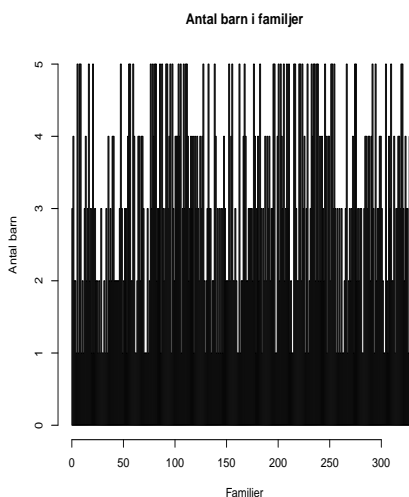


Figur 4: Sjukdomsspårning

För att kunna skatta hur många barn i skolan som bär sjukdomen och hur vi kommer åt familjer som har sjuka barn, använder vi simuleringar. Vi bestämmer först ett barnantal och sedan simulerar vi vilka av barnen som bär sjukdomen med sannolikhet p_s . Efter det delar vi slumpmässigt upp barnen i olika stora familjer. Familjerna är likformigt fördelade med mellan 1 och 5 barn, $X_i \sim U(1, 5)$. I diagrammen nedan ser vi hur de stokastiska familjerna ser ut. Vi bestämmer oss för att studera de två fallen då vi har 100 respektive 1000 barn. Observera alltså att det är slumpen samt val av fördelning för familjernas storlekar som avgör hur många familjer som bildas. I de fall vi betraktar delas 100 barn upp i 39 familjer och 1000 barn upp i 335 familjer.



Figur 5: 100 barn, 39 familjer



Figur 6: 1000 barn, 335 familjer

Nedan presenterar vi simuleringsresultatet då vi slumpmässigt har valt ett barn respektive en familj. Vi har upprepat processen 10 gånger med återläggning. Vid varje slumpmässigt val har alla barn respektive alla familjer lika stor sannolikhet att bli utvald, och det är alltså möjligt att ett och samma barn eller en och samma familj blir vald flera gånger. I tabellerna 5-8 står slumpfallena i första kolumnen, i andra kolumnen står numret för ett slumpmässigt valt barn eller en slumpmässigt vald familj och i tredje kolumnen står barnantalet i det utvalda barnets familj respektive antalet barn i den utvalda familjen. De fyra sista kolumnerna beskriver hur många sjuka barn som finns i den utvalda familjen, då sannolikheten för sjukdomen varierar.

Tabellerna 5-8 visar att sjukdomsspårningen ger olika resultat om vi väljer en familj eller om vi väljer ett barn i skolan slumpmässigt. Då vi väljer ett barn slumpmässigt, så är sannolikheten större att barnet kommer från en stor familj (4-5 barn). Detta beror på urvalsmetoden, ty sannolikheten att välja ett barn från en viss familj är proportionellt mot storleken av familjen. Därav följer att vi i denna situation lättare kan spåra sjukdomen: dels har vi flera barn att observera (ett slumpmässigt barn och dennes flera syskon), dels är det större sannolikhet att något av barnen i en stor familj bär sjukdomen.

Om vi däremot väljer en familj slumpmässigt, så är det mindre sannolikt att det sker en belöning och att vi hittar ett barn som bär sjukdomen. Denna urvalsmetod är således mindre effektiv för att spåra sjuka barn i skolan.

Då sannolikheten för sjukdomen blir liten blir det naturligtvis svårare att spåra sjuka barn. Vi ser i tabellerna nedan att man kan då inte märka speciellt stor skillnad mellan belöningar vid de två olika urvalsmetoderna.

Slumptillfälle	Slumpmässigt valt barn					
	j	X_n	$p_s = 0.5$	$p_s = 0.25$	$p_s = 0.10$	$p_s = 0.05$
1	51	4	2	1	0	0
2	44	5	2	1	1	0
3	58	5	1	1	0	0
4	93	4	3	0	0	0
5	57	5	1	1	1	0
6	55	5	1	1	0	0
7	92	4	3	0	0	0
8	96	2	2	0	0	1
9	38	3	2	1	1	0
10	50	2	1	1	0	0

Tabell 5: Sjukdomsspårning, slumpmässigt valt barn (100 barn, 39 familjer).

Slumptillfälle	Slumpmässigt vald familj					
	n	X_n	$p_s = 0.5$	$p_s = 0.25$	$p_s = 0.10$	$p_s = 0.05$
1	28	1	1	0	0	0
2	4	1	1	0	0	0
3	26	1	0	0	0	0
4	25	1	1	0	1	0
5	26	1	0	0	0	0
6	30	2	0	1	0	0
7	13	4	3	1	0	0
8	22	5	1	1	0	0
9	23	1	1	0	0	0
10	6	5	1	1	0	0

Tabell 6: Sjukdomsspårning, slumpmässigt vald familj (100 barn, 39 familjer).

Slumptillfälle	Slumpmässigt valt barn					
	j	X_n	$p_s = 0.5$	$p_s = 0.25$	$p_s = 0.10$	$p_s = 0.05$
1	268	5	2	2	0	0
2	575	5	5	2	1	0
3	860	4	3	3	0	0
4	344	3	1	0	0	0
5	513	4	1	1	1	1
6	356	3	1	0	0	0
7	612	3	0	0	0	0
8	807	2	2	1	0	1
9	345	3	1	0	0	0
10	189	1	0	0	0	0

Tabell 7: Sjukdomsspårning, slumpmässigt valt barn (1000 barn, 335 familjer).

Slumptillfälle	Slumpmässigt vald familj					
	n	X_n	$p_s = 0.5$	$p_s = 0.25$	$p_s = 0.10$	$p_s = 0.05$
1	115	3	2	1	0	0
2	327	1	1	0	0	0
3	88	5	3	1	1	0
4	31	1	0	0	0	0
5	168	5	4	3	0	1
6	291	4	0	1	0	0
7	103	3	1	0	0	0
8	181	1	0	0	0	0
9	3	2	1	1	0	0
10	201	1	0	0	0	0

Tabell 8: Sjukdomsspårning, slumpmässigt vald familj (1000 barn, 335 familjer).

3.5 Köer

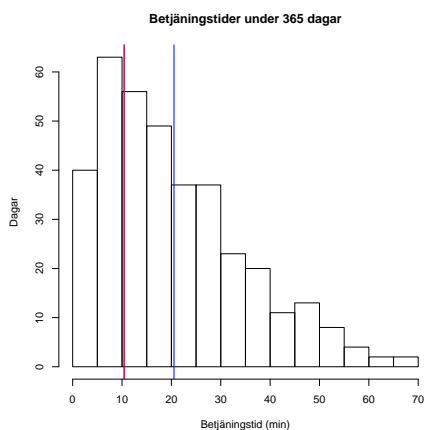
3.5.1 Alltid sysselsätt system

Vi tänker oss en kö med en betjäningstation. Det finns alltid minst en person i kön. Systemet är alltså aldrig tomt. Han som betjänar kunder arbetar 6 timmars arbetspass. Betjäningstiderna, X_i , är oberoende av varandra och har samma fördelning. Betjäningstationen är öppet alla dagar om året (365 dagar). Varje dag under ett år observerar vi hur lång betjäningstiden är för den betjäning som pågår då arbetspasset har pågått i fem timmar. Vi simulerar alltså betjäningstider för hela arbetspassen under samtliga 365 dagar och kontrollerar hur lång servicetiden är då personen har arbetat fem timmar. Vi studerar två olika fall med olika fördelningar men med samma väntevärde, 10 minuter, för betjäningen.

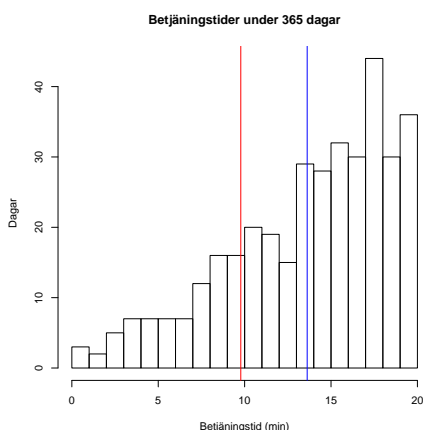
Vi antar först att betjäningstiderna är exponentialfördelade med intensitet $\frac{1}{10}$. Under de 365 dagarna är medelvärdet för betjäningstiden för den betjäning som pågår efter 5 timmars arbete 20 minuter och 30 sekunder, vilket är mer än dubbla väntevärdet. Vi ser också att betjäningstiden under 262 dagar har överstigit väntevärdet (se Figur 7).

Vi justerar vårt antagande och antar nu istället att betjäningstiderna är likformigt fördelade mellan 0 och 20 minuter. I detta fall är medelvärdet för betjäningstiden för den betjäning som pågår efter 5 timmars arbete 13 minuter och 40 sekunder, och betjäningstiderna har överstigit väntevärdet under 283 dagar (se Figur 8).

Anmärkning 3.1. Att betjäningstiderna vid vår valda tidpunkt skiljer sig så mycket från väntevärdet beror på att långa betjäningstider blir överrepresentativa och får en stor vikt då vi beräknar medelvärdet under 365 dagar.



Figur 7: $X_i \sim \exp(\frac{1}{10})$



Figur 8: $X_i \sim U(0,20)$

3.5.2 Väntetid i $M/G/1$ -kö

Vi kan tänka sig mer komplicerade tillämpningar om vi tar hänsyn till ködicipliner. Vi antar att vi har en $M/G/1$ -kö, se kapitel 2.3, där kunder anländer enligt en poissonprocess med ankomstintensitet λ . För $i \geq 1$ och $j \in \{1, \dots, 100\}$ betecknar vi ankomsttiden för kund nummer i dag j med $T_{i,j}$, väntetiden i kö för kund nummer i dag j med $W_{i,j}$ och betjäningstiden för kund nummer i dag j med $S_{i,j}$. Ankomsttiderna och betjäningstiderna för alla kunder är lika fördelade och oberoende av varandra. Vi antar att vi kommer till ett kösystem under 100 dagar, varje dag vid tidpunkten t . Vi antar att vi dag j är kund nummer n_j . Vi är intresserade av att studera hur stor skillnad som uppstår i väntetid i kö mellan den n_j :te kunden och kunderna före honom.

För varje dag j simulerar vi en kö med samma ankomstintensitet och med samma fördelning för betjäningstiderna. Talet n_j är slumpmässigt och varierar för dag till dag och beror naturligtvis på de $n_j - 1$ första kundernas ankomsttider och den gemensamma ankomstintensiteten. Efter att ha simulerat ankomsttider och betjäningstider för respektive kund beräknar vi motsvarande väntetider. Vi observerar varje dag hur länge den n_j :te kunden får vänta i kön och jämför detta med medelväntetiden i kö för de $n_j - 1$ första kunderna. Slutligen beräknar vi medelväntetid i kö för den n_j :te kunden under 100 dagar, vilket vi jämför med medelväntetid för övriga kunder som har kommit till systemet före tidpunkten t under de 100 dagarna.

Väntetiden i kö för den första kunden är naturligtvis noll, ty systemet är alltid tomt när första kunden anländer. Andra kundens väntetid i kö är noll eller högst första kundens betjäningstid. Kunder därefter anländer antingen till ett tomt system eller då systemet är upptaget av en eller flera kunder. Om det finns en kund i systemet, då den i :te kunden anländer, så får han vänta tills den $(i - 1)$:te kunden betjänats klart. Vi har följande samband (oberoende av dag)

för den i :te kunden, då $i \geq 2$:

$$W_{i,j} = \max(W_{(i-1),j} + S_{(i-1),j} - (T_{i,j} - T_{(i-1),j}), 0), \quad (17)$$

$$W_{i,j} + S_{i,j} = \text{tid i systemet för den } i\text{:te kunden}, \quad (18)$$

$$T_{i,j} + W_{i,j} + S_{i,j} = \text{tidpunkten då den } i\text{:te kunden lämnar systemet}. \quad (19)$$

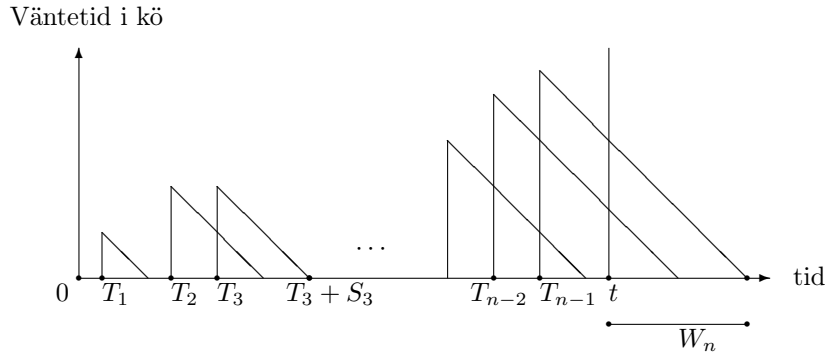
Vi använder ekvation (17) samt informationen att första kundens väntetid i kö är noll och andra kundens kötid högst första kundens betjäningstid för att beräkna medelväntetiden i kö, \bar{W}_j , för de $(n_j - 1)$ första kunderna dag j :

$$\bar{W}_j = \frac{\max(T_{1,j} + S_{1,j} - T_{2,j}, 0)}{n_j - 1} + \frac{\sum_{i=3}^{n_j-1} \max(W_{(i-1),j} + S_{(i-1),j} - (T_{i,j} - T_{(i-1),j}), 0)}{n_j - 1}. \quad (20)$$

Väntetiden i kö för den n_j :te kunden ges av

$$W_{n_j} = \max(W_{(n_j-1),j} + S_{(n_j-1),j} - (t - T_{(n_j-1),j}), 0). \quad (21)$$

Följande bild illustrerar köprocessen:



Figur 9: M/G/1-kö

Vi presenterar nu tre konkreta räkningar, där $t = 120$ minuter, som visar att om vi anländer till kösystemet vid en viss tidpunkt t under 100 dagar så kommer vi ha längre väntetid i kö än medelväntetiden för kunderna innan oss.

Vi varierar ankomstintensiteten λ men behåller fördelningen för betjäningstiderna, $S_i \sim U(1, 3)$. Väntevärdet för betjäningstiden $E[S]$ är 2 minuter.

I första fallet har vi att $\lambda = 0.25$ så kunderna kommer till systemet enligt en poissonprocess var fjärde minut, $T_i \sim \exp(\frac{1}{4})$. Då den n_j :te kunden anländer, så har det kommit och gått ett slumpmässigt antal kunder till och från systemet. För att ge läsaren en bild hur köprocessen kan se ut under en dag, så visar vi simuleringsresultatet för dag 100. I tabell 9 ser vi till exempel att $n_{100} = 35$ och

att väntetid för den n_{100} :te kunden är lika med noll. Observera dock att tabell 9 endast visar tider för en dag, och vi är intresserade av totalt 100 dagar.

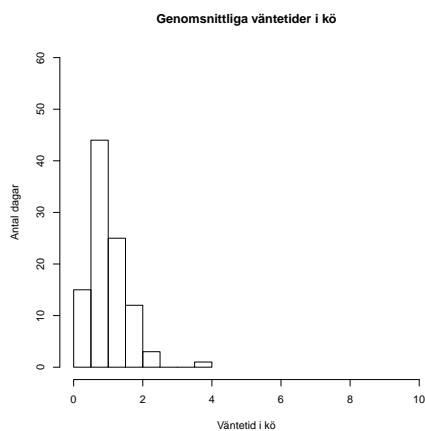
Kund	Ankomsttid	Väntetid i kö	Betjäningstid	Avgångstid
1	6.904267	0.000000	2.061220	8.965487
2	12.534594	0.000000	2.493027	15.027621
3	14.171789	0.8558317	1.940204	16.967825
4	21.484120	0.000000	2.280522	23.764643
5	23.885748	0.000000	1.905098	25.790846
6	24.338247	1.4525986	1.733799	27.524644
7	26.903776	0.6208685	1.166829	28.691473
8	29.271126	0.000000	1.840912	31.112038
9	38.264853	0.000000	2.408973	40.673826
10	39.071700	1.6021258	2.237984	42.911810
11	41.288646	1.6231639	1.309656	44.221466
12	41.529205	2.6922613	2.168514	46.389980
13	43.067462	3.3225178	1.861059	48.251039
14	44.252069	3.9989699	2.733575	50.984614
15	46.562808	4.4218057	2.177120	53.161734
16	50.714400	2.4473337	1.681144	54.842878
17	51.632026	3.2108520	2.164931	57.007809
18	52.655445	4.3523642	2.995267	60.003077
19	53.997618	6.0054583	1.104707	61.107784
20	62.097954	0.000000	2.217866	64.315820
21	67.031228	0.000000	1.624036	68.655264
22	69.913741	0.000000	2.173349	72.087090
23	70.093557	1.9935324	1.905388	73.992478
24	74.179611	0.000000	1.249259	75.428871
25	76.885110	0.000000	1.806434	78.691544
26	78.256185	0.4353590	2.025114	80.716658
27	78.319408	2.3972495	2.324518	83.041176
28	79.767667	3.2735090	2.186169	85.227345
29	88.809364	0.000000	2.945772	91.755136
30	103.082613	0.000000	2.035306	105.117919
31	103.605589	1.5123307	2.924429	108.042348
32	108.905085	0.000000	2.971619	111.876705
33	109.148142	2.7285628	1.943092	113.819797
34	116.726608	0.000000	1.054139	117.780747
35	120.000000	0.000000	1.249273	121.249273

Tabell 9: Köprocessen för dag 100, $\lambda=0.25$.

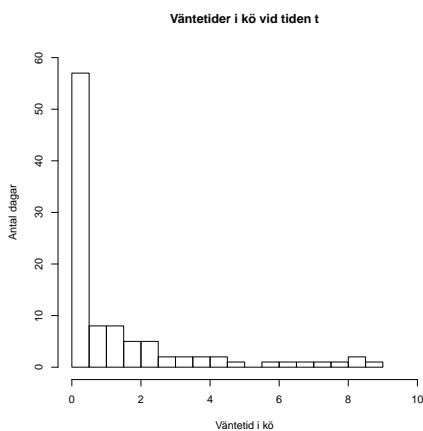
Medelväntetiden i kön för de $n_j - 1$ första kunderna under 100 dagar blir $\sum_{j=1}^{100} \frac{\bar{W}_j}{100} = 1.01$ minuter och den genomsnittliga väntetiden i kö då vi anländer efter 2 timmar blir $\sum_{j=1}^{100} \frac{W_{n_j}}{100} = 1.31$ minuter. Under de 100 observerade

dagarna får vi alltså vänta totalt 0.3 minuter längre per dag än övriga kunder i genomsnitt, och $W_{n_j} > \bar{W}_j$ vid 36 tillfällen.

Vi illustrerar våra simuleringar i diagram nedan. I diagrammet till vänster ser vi genomsnittliga väntetider i kö, \bar{W}_j , och på högersidan ser vi väntetider i kö, W_{n_j} , för den bestämda tidpunkten t . Oavsett att den n_j :te kundens väntetid i kö har varit noll under flera dagar enligt Figur 11, så ger "långa svansar" en stor vikt, och medelväntetiden i kö blir genomsnittligt längre för den n_j :te kunden än för de kunder som har besökt systemet före honom.

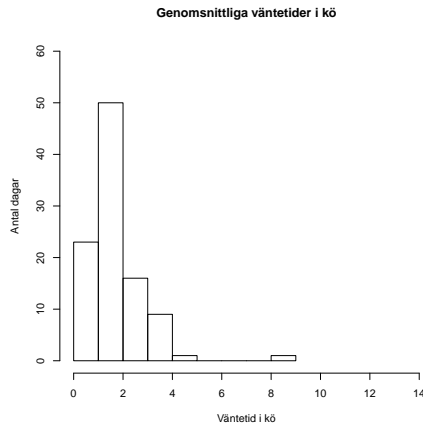


Figur 10: $\lambda = 0.25$

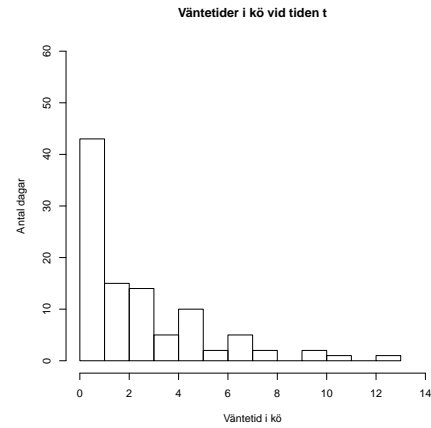


Figur 11: $t = 120$ min, $\lambda = 0.25$

I det andra fallet har vi samma antaganden som tidigare, men tar i stället $\lambda = 0.33$ så en kund kommer alltså ungefär var tredje minut, $T_i \sim \exp(\frac{1}{3})$. Under 100 dagar är den genomsnittliga väntetiden i kön för de $n_j - 1$ första kunderna 1.69 minuter och den genomsnittliga väntetiden då vi anländer 2.29 minuter. Under de observerade dagarna får vi vänta i kö 0.6 minuter längre per dag än övriga kunder i genomsnitt, och $W_{n_j} > \bar{W}_j$ vid 48 tillfällen.

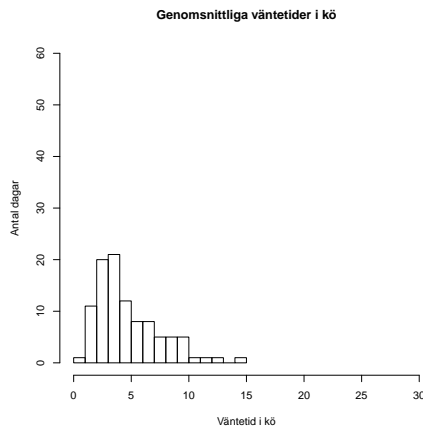


Figur 12: $\lambda = 0.33$

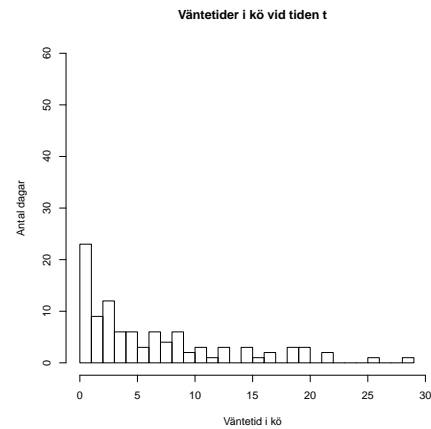


Figur 13: $t = 120$ min, $\lambda = 0.33$

I det tredje fallet har vi samma antaganden som ovan, men låter $\lambda = 0.45$, $T_i \sim \exp(\frac{1}{2.22})$, varav följer att $\lambda E[S] < 1$. Det är naturligt att det nu oftare bildas längre köer än i tidigare fall. Under 100 dagar är den genomsnittliga väntetiden i kön 4.66 minuter och den genomsnittliga väntetiden då vi anländer 6.35 minuter. Under de observerade dagarna får vi vänta totalt 1.69 minuter längre per dag än övriga kunder i genomsnitt, och $W_{n_j} > \bar{W}_j$ vid 55 tillfällen.



Figur 14: $\lambda = 0.45$



Figur 15: $t = 120$ min, $\lambda = 0.45$

Vi kan konstatera att i exemplen ovan har vi längre väntetid då vi anländer vid tidpunkten t än medelväntetiden för kunderna innan oss. Detta visar sig gälla även fast ankomstintensiteten är låg ($\lambda = 0.25$) och fördelningen för betjäningstiderna är oförändrad. Att vår väntetid i kö blir längre beror på obser-

vationsplanen och length-biased sampling: vi har större sannolikhet att komma till kön då det pågår en lång betjäning än att komma under en kort betjäning. Att vi har större sannolikhet att komma under en lång betjäning medför också att sannolikheten för en lång väntetid i kön blir större.

4 Diskussion

Inspektionsparadoxen uppstår inom förnyelseteorin och finns inbyggd i många vardagliga situationer. Det är av stort intresse att inspektionsparadoxen även naturligt uppstår i diskreta fall.

I uppsatsen har vi försökt göra verklighetstroga simuleringar för att illustrera inspektionsparadoxen numeriskt. Naturligtvis inser man att vid annorlunda observationsplaner, som till exempel med andra antaganden och andra val av fördelningar, kan resultaten se annorlunda ut. Vi anser att simulering har varit en lämplig metod att skapa datamaterial till uppsatsen. Då vi studerar inspektionsparadoxen utgår vi ofta från att vi har godtyckligt många förnyelseintervall och att vi ska välja en tidpunkt som ligger långt bort på tidsaxeln. Detta kommer man effektivt åt genom att göra simuleringar.

Vi är medvetna om att vår studie av inspektionsparadoxen är begränsad i den mening att vi endast studerat ett givet antal simulerade exempel. Som en fortsättning på denna uppsats vore det till exempel intressant att studera om det finns flera sätt att tillämpa köteorin än de Cooper med flera [2] presenterar.

Referenser

- [1] J. E. Angus. An asymptotic expansion for the inspection paradox. *Probability in the Engineering and Informational Sciences*, 20:87–94, 2006.
- [2] R. B. Cooper, S-C. Niu, and M. M. Srinivasan. Some reflections on the renewal-theory paradox in queueing theory. *Journal of Applied Mathematics and Stochastic Analysis*, 11(3):355–368, 1998.
- [3] Stein W. E. and Dattero R. Sampling bias and the inspection paradox. *Mathematics Magazine*, 58(2):96–99, March 1985.
- [4] D. Hemenway. Why your classes are larger than “average”. *Mathematics Magazine*, 55(3):162–164, May 1982.
- [5] W. Herff, U. Jochems, and U. Kamps. The inspection paradox with random time. *Statistical Papers*, 296(38):103–110, 1997.
- [6] S. M. Ross. The inspection paradox. *Probability in the Engineering and Informational Sciences*, 17:47–51, 2003.
- [7] S. M. Ross. *Introduction to Probability Models*. Elsevier/Academic Press, Amsterdam, ninth edition, 2007. International edition.
- [8] H. L. Royden. *Real analysis*. Prentice Hall, third edition, 1988.
- [9] Hu T. and Zhuang W. Likelihood ratio ordering of the inspection paradox. *Probability in the Engineering and Informational Sciences*, (18):503–510, 2004.
- [10] J. D. Terwilliger, W. D. Shannon, G. M. Lathrop, J. P. Nolan, L. R. Goldin, G. A. Chase, and D. E. Weeks. True and false positive peaks in genomewide scans: Applications of length-biased sampling to linkage mapping. *The American Society of Human Genetics*, 61:430–438, May 1997.
- [11] J. Virtamo. Lecture notes from a course in Queueing Theory given at Hensinki University of Technology. <http://www.netlab.tkk.fi/opetus/s38143/luennot/english.shtml>.
- [12] J. R. Wilson. The inspection paradox in renewal-reward processes. *Operations Research Letters*, 2(1):27–30, April 1983.