

Extra Uppgifter II SDA1

Mona Sfaxi och Mattias Villani

February 12, 2024

Dessa övningar bör lösas för hand med hjälp av en miniräknare och kursens formelblad.

Övningar markerade med en * kan lösas av dem som vill veta lite mer kring olika detaljer, men de kommer inte att vara så relevanta för tentamen.

Uppgift 1

På en arbetsplats är sannolikheten att en slumpmässigt vald person dricker kaffe 0.80 och för te motsvarar denna siffra 0.55. Sannolikheten att man varken dricker te eller kaffe är 0.12. Vad är sannolikheten att man antingen dricker te eller kaffe? Vad är sannolikheten att man dricker både te och kaffe? Går det att rita dessa sannolikheter i ett Venndiagram?

Uppgift 2

Thorin funderar på att besöka sin vän Bilbo. Han har 12 vänner som han skulle kunna tänka sig ha som sällskap dit.

a.)

Han tänkte ringa 4 av dem, på hur många olika sätt kan han göra detta på?

b.)

Vid närmare eftertanke så insåg Thorin att han faktiskt ville ha med sig alla 12 vänner istället för 4, så han bestämde sig för att ringa alla 12. På hur många olika sätt kan han göra detta?

c.)

Det visade sig att alla hans vänner ville följa med men tyvärr har Thorin bara plats för 4 stycken vänner i sin bil. Han väljer slumpmässig 4 av dem som ska åka i hans bil och för honom spelar det ingen roll var de sitter någonstans. På hur många olika sätt kan han göra detta?

Uppgift 3.)

Låt händelsen:

$A = \{\text{Har ett volontärsarbete}\}$, och $A^c = \{\text{Har inte ett volontärsarbete}\}$. Det visar sig att i Stockholm så är $P(A) = 0.35$. Vidare så kan vi låta:

$B = \{\text{Gillar Tacos}\}$ och $B^c = \{\text{Gillar inte Tacos}\}$. Sedan vet vi även att $P(B^c) = 0.67$ och även att sannolikheten att man gillar Tacos och har ett volontärsarbete är 0.25. Nedan har vi en korstabell som ska illustrerar de simultana sannolikheterna men någon har glömt att fylla i den.

Händelse	A	A^c	Marginal
B			
B^c			
Marginal			1

a.)

Hjälp till att fylla i tabellen.

b.)

Vad är sannolikheten att åtminstone någon av händelserna A eller B inträffar?

c.)

Vad är sannolikheten att åtminstone någon av händelserna B och A^c inträffar?

d.)

Tror du att det är mer troligt att gilla Tacos om man har ett volontärsarbete jämfört med om man inte har ett volontärsarbete?

Uppgift 4

Kusinerna Diogenes och Alexander brukar varje lördagskväll träffas för att spela i ett gamingcenter. De väljer då mellan att spela två olika spel; Zelda eller Super Mario. Sannolikheten att de spelar Zelda och att Diogenes anländer först är 0.2448. Vidare vet vi att om de spelar Zelda är sannolikheten 0.4766355 att det var Diogenes som anlände först. Slutligen vet vi även att om Alexander anländer först är sannolikheten 0.42 att de spelar Zelda. Finn sannolikheten att Alexander anländer först (Tips: använd Bayes sats och lagen om total sannolikhet).

Uppgift 5

Palmyra är ofta väldigt tankespridd och glömmer ibland sin matlåda hemma ett visst antal gånger i veckan. Nedan ser vi fördelningen över antal gånger hon glömmer dem/vecka. Där y betecknar händelsen att hon glömmer den hemma och $p(y)$ betecknar sannolikheten:

y	0	1	2	3	4	5
$p(y)$	0.09	0.15	0.27	0.25	0.10	0.14

a.)

Beräkna väntevärdet för Y .

b.)

Beräkna variansen och standardavvikelsen för Y .

c.)

Glömmer hon matlådan hemma så finns det som tur är en restaurang vid hennes jobb som hon kan köpa mat ifrån. En lunch där kostar 120kr. Hur mycket behöver hon i genomsnitt betala per vecka?

Uppgift 6

Kurt har en liten glassfristad på sitt sommarställe där han tillverkar glasstrutar och säljer till sina grannar. Han har två smaker; Jordnöt med salt-karamellsås doppad i choklad och smaken Brownie med vanilla swirl. Jordnötsglassen säljs i två-pack och Brownie-glassen säljs styckvis. Nedan syns den simultana fördelningen över hur mycket han säljer per dag:

	Jordnöt			Marginal Brownie
Brownie	0	2	4	
1	0.08	0.00	0.01	0.09
2	0.12	0.17	0.06	0.35
3	0.36	0.16	0.04	0.56
Marginal Jordnöt	0.56	0.33	0.11	1

a.)

Beräkna väntevärdet för båda glassmakerna.

b.)

Finns det något samband mellan försäljningen för de båda glassmakerna, beräkna kovariansen för dem.

Uppgift 7

Vi har tre slumpvariabler X , Y och Z som har väntevärden 15, 11 och 18. Nedan syns även deras kovariansmatris:

$$\begin{matrix} & X & Y & Z \\ \begin{matrix} X \\ Y \\ Z \end{matrix} & \begin{bmatrix} 25 & -7 & 13 \\ -7 & 9 & 8 \\ 13 & 8 & 20 \end{bmatrix} \end{matrix}$$

a.)

Beräkna Väntevärdet och Variansen för $3X + 5Y$

b.)

Beräkna väntevärdet och variansen för $4X - 5Y + 7$

c.)

Beräkna väntevärdet och standardavvikelsen för $2X + 5Y + 3Z$

d.)

Vilka två variabler har den starkaste korrelationen?

Uppgift 8

Adolfsbergsskolan annordnar varje år en tävling i friidrott. I en av grenarna längdhopp vet man att längden på hoppen bland 10 åriga flickor är normalfördelad med väntevärde 2.43 och en varians på 0.2025.

a.)

Vad är sannolikheten att man hoppar under 2.8 meter?

b.)

Vad är sannolikheten att man hoppar under 2 meter?

c.)

Hur stor är sannolikheten att man hoppar mellan 2.2 och 2.6 meter?

d.)

Vad är sannolikheten att man hoppar över 3 meter?

e.)

Julia och Angelica var de två sista flickorna som hoppade. En lärare hade antecknat allas resultat. Tyvärr hade han tappat bort resultaten för de sista två hoppen. Lyckligtvis hade han antecknat sannolikheterna för dem på ett annat papper och kommit fram till att sannolikheten att man minst hoppade så långt som Julia var 0.075 och att sannolikheten att man högst hoppade så långt som Angelica var 0.33. Hur långt hoppade båda flickor?

Uppgift 9

Efter att Alexander och Diogenes har spelat hela natten brukar de gå hem till Alexanders mamma och äta söndagslunch. Maten brukar bli klar någon gång mellan klockan 11:00-13:00, men det är lika troligt att maten exempelvis blir klar den första halvtimmen som den sista halvtimmen.

a.)

Vilken fördelning kan vi modellera detta med? Motivera, och rita fördelningen samt sätt ut gränserna.

b.)

Vad är sannolikheten att maten blir klar mellan 11 och 12? Rita ut detta område i en graf.

c.)

Vad är sannolikheten att maten blir klar mellan 11:30 och 12:45? Rita ut och markera detta område i en graf. Är detta en area eller en sannolikhet?

d.)

När kan maten förväntas vara klar?

Uppgift 10

a.)

Hälsotidningen *Collosseum*, gjorde en undersökning där de frågade 200 slumpmässigt utvalda personer ifall de styrketränar minst 1 gång i månaden och 54 stycken svarade ja.

a.)

Skatta proportionen som styrketränar samt standardfelet.

b.)

Konstruera ett 95%-igt konfidensintervall för proportionen som styrketränar och tolka intervallet.

Uppgift 11*¹

Antag att att en slumpvariabel X är normalfördelad med det sanna men okända väntevärdet μ men att standardavvikelsen σ är känd. Antag vidare att man drar ett stickprov av storlek n för att skatta det sanna väntevärdet och för att konstruera ett konfidensintervall för det. Härled den lägre och övre gränsen av ett 95%-igt konfidensintervall genom att använda:

$$P\left(-Z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}\right) = 1 - \alpha$$

Uppgift 12

Utifrån utskriften på övning 23.1 i boken SDM, Beräkna och tolka ett 95% prediktionsintervall för Earnings för en person med testpoängen $SAT = 1350$ poäng, vilket är nära 90% percentilen för alla som gör testet: [\[Länk\]](#). Antag att medelvärdet för SAT i stickprovet är 1075.

¹Kommer ej på tentan men kan vara bra för förståelsen kring konfidensintervall

Uppgift 13

a.)

Enligt en undersökning som gjorts på beställning av ett stort företag som säljer kaffe dricker 80% av Sveriges befolkning kaffe regelbundet. Studenten Ayla ville testa om detta stämde och ställde sig därför på SU:s campus och frågade 120 personer ifall de drack kaffe regelbundet (hon definierade regelbundet som minst 2 gånger i veckan) och 104 personer svarade ja.

b.)

Testa hennes hypotes och beräkna p-värdet för det. Kan man förkasta H_0 på en 5%-ig signifikansnivå?

c.)

Ayla såg resultatet från testet och bestämde sig för att istället göra ett enkelsidigt test, där hon istället var intresserad av att testa ifall andelen som dricker kaffe var större än 0.8. Ställ upp hypoteserna för det nya testet, beräkna teststatistikan och därefter p-värdet. Diskutera problemet med att ställa upp nya hypoteser efter att redan ha gjort studien och utfört testet på fråga b.). Hade man kunnat redovisa resultaten på ett annat sätt än genom att använda sig av hypotestest? Diskutera!

d.)

Nämn vilka antaganden man behöver ha för att kunna utföra testet i b.) samt om dessa är uppfyllda. Diskutera även om det finns något problem med Aylas tillvägagångssätt för att samla in sitt datamaterial och om resultatet är tillförlitligt.

Uppgift 14

a.)

Diego hörde av en vän att en genomsnittlig puls låg på ca 73 (per minut) för en person i hans ålder. Han har inte tränat regelbundet på ett långt tag och misstänker att hans genomsnittliga puls ligger över medelvärdet. För att bedriva tiden en timme innan han hade föreläsning gjorde han 5 stycken mätningar av sin puls på måfå under denna timme och fick värdena:

Diegos puls	75	74	73	74	73
-------------	----	----	----	----	----

Testa hans hypotes på en 10%-ig signifikansnivå. Ställ tydligt upp dina hypoteser, teststatistikan under H_0 , kritiskt värde och vilken slutsats man kan dra. Måste han utgå från några antaganden innan han utför testet? Motivera

b.)

När han ändå höll på så blev han fundersam ifall hans genomsnittliga puls blir påverkad av föreläsningarna så efter föreläsningen gjorde han också 5 stycken mätningar av sin puls. Tabellen nedan sammanfattar resultatet:

Före föreläsningen	75	74	73	74	73
Efter föreläsningen	77	75	76	74	75

Testa ifall hans genomsnittliga puls skiljer sig åt innan och efter föreläsningen på en 5%-ig signifikansnivå. Börja med att nämna vilka antaganden du måste göra. Utför sedan testet genom att ställa upp hypoteser, kritiskt värde och vilken slutsats man kan dra.

c.)

Skriv ner R-kommandot som beräknar p-värdet för ditt test på fråga b.).

Uppgift 15

Vännerna Diego och Ernesto argumenterade med varandra om vilken ort som var kallast under vintern. Diego som bor i Västerås hävdade strängt att hans stad var kallast medan Ernesto ansåg att Stockholm var kallast. Då Ernesto tyckte att SMHI och andra vädersidor var opålitliga så bestämde de sig för att mäta temperaturen själva i sina hemstäder under 5 slumpmässigt utvalda dagar i December-Februari under vintern 2022-2023. De kom fram till följande resultat:

Datum	7/12	17/12	28/1	12/2	24/2
Västerås	-8.2	-4.4	0.7	-1	-2.7
Stockholm	-4.7	-3.4	0.8	0.9	-0.5

Testa ifall Diego hade rätt på en 1%-ig signifikansnivå, börja med att skriva upp dina antaganden, om du tycker att de är rimliga samt om detta är ett parat eller icke-parat test och utför sedan testet och dra en slutsats. Vilken är den lägsta signifikansnivån man kan förkasta nollhypotesen på?