UPPSALA UNIVERSITET
Matematiska Institutionen
Annika Heckel

PROV I MATEMATIK
Sannolikhet och statistik DV, 1MS321
2022-08-16

Hjälpmedel: Räknedosa, formel- och tabellsamling och engelsk-svensk ordlista för kursen 1MS321. Goda resonemang och motiveringar vägs in vid poängsättning. För betygen 3, 4, resp. 5 krävs normalt minst 18, 25, resp. 32 poäng.

Fråga 1:

a) Låt X vara en diskret slumpvariabel som endast har utfallen 0, 1 och 2 med positiv sannolikhet. Det är givet att $P(X = 0) = P(X = 2) = 0.2$. Beräkna standardavvikelsen av X. (2p)

b) Låt Y vara en slumpvariabel med täthetsfunktion

$$f(y) = \frac{1}{2}y \qquad \text{när } 0 \leqslant y \leqslant 2$$

och $f(y) = 0$ annars. Bestäm $P(Y < 1)$ och $E[Y]$. (2p)

c) Beräkna kovariansen mellan slumpvariablerna Y och $1/Y$. (2p)

Fråga 2: Vi vet att 1% av befolkningen har en särskild sjukdom, och att det finns ett test för denna sjukdom. En sjuk person som testar sig får ett positivt utslag med 95% sannolikhet. Om en frisk person tar testet ger det ett negativt utslag med 99% sannolikhet.

a) Du väljer en person på måfå ur befolkningen, och testar denna. Vad är sannolikheten att testet ger positivt utslag? (1p)

b) Återigen, du väljer en person på måfå ur befolkningen, och testar denna. Testet ger ett positivt utslag, vad är sannolikheten att hen har sjukdomen? (2p)

c) Hur stor andel av befolkningen behöver vara sjuk för att sannolikheten i b) är minst 90%? Med andra ord, hur vanligt förekommande behöver sjukdomen vara för att, en slumpmässigt vald testad person med positivt utslag, är sjuk med 90% sannolikhet? (2p)

Fråga 3: På ett visst socialt medium kan användare ladda upp videoklipp. Hur storleken av filerna är fördelade är okänd, men företaget bakom mediet vet att en genomsnittlig fil har storleken 500MB och att standardavvikelsen är 300MB.

Antag att 150 användare laddar upp ett klipp var, oberoende av varandra. Uppskatta sannolikheten att 80GB är tillräckligt med utrymme för att förvara all dessa användares videoklipp. Introducera en lämplig modell och definiera alla slumpvariabler i modellen.

(6p)

Fråga 4: För ett heltal $n \geqslant 1$, låt $x_1, \ldots, x_n$ vara ett oberoende stickprov av slumpvariabeln $X \sim N(\mu, 2\sigma^2)$, där μ och $\sigma > 0$ är okända parametrar. Oberoende av detta, låt y vara ett stickprov av slumpvariabeln $Y \sim N(\mu, \sigma^2)$, där μ och σ är samma okända parametrar. Låt $\bar{x}$ stå för stickprovsmedelvärdet av $x_1, \ldots, x_n$.

a) Visa att både $\hat{\mu}_1 = \frac{3\bar{x}+y}{4}$ och $\hat{\mu}_2 = \frac{\bar{x}+3y}{4}$ är väntevärdesriktiga skattningar av μ.

(2p)

b) För vilka värden av n är $\hat{\mu}_1$ en effektivare skattning än $\hat{\mu}_2$? (3p)

Fråga 5: En leverantör av en plattform för e-post vill markera vissa e-post som viktiga. En grupp ingenjörer designar en algorithm och påstår att den markerar 15% av all inkommande e-post som viktig. Som ett test skickar företaget 200 e-post till en adress. Av dessa markeras 44 e-post som viktiga.

a) Ange ett ungefärligt 95% konfidensintervall för sannolikheten att en e-post ska markeras som viktig. Motivera ditt svar ordentligt. Vad kan du säga om ingenjörernas påstående? (4p)

b) Vad blir det ungefärliga konfidensintervallet och din slutsats kring ingenjörernas påstående om konfidensnivån är 99% istället? (2p)

Fråga 6: Givet en Markovkedja $X_0, X_1, X_2, \dots$ med tillståndsrum $E = \{0, 1, 2\}$. Markovkedjan förflyttas från tillstånd 0 till tillstånd 1 med sannolikheten $\frac{1}{4}$. Vidare, markovkedjan stannar kvar på tillstånd 1, med sannolikheten $\frac{1}{2}$. Om kedjan befinner sig på tillstånd 2 kan den inte stanna kvar, men måste flytta till ett annat tillstånd. Oavsett vilket tillstånd Markovkedjan befinner sig i, är sannolikheten alltid precis $\frac{1}{2}$ att det efterkommande tillståndet är 0.

a) Bestäm övergångsmatrisen $\mathbf{P}$ för Markovkedjan. (2p)

b) Skissa Markovkedjans övergångsdiagram. (1p)

c) Markovkedjan börjar vid tidpunkt $t = 0$ med tillståndet 0. Beräkna $P(X_2 = 1)$. (1p)

d) Bestäm Markovkedjans stationära fördelning. (Du kan anta att en unik stationär fördelning existerar.) (2p)

Fråga 7: Du får notifikationer för e-post och sms i din telefon. Du tar emot genomsnittligen 0,5 e-post och 2 sms per timma och fördelningen för antalet e-post och sms över tiden kan antas vara fördelade efter två oberoende Poissonprocesser. Introducera en lämplig modell, tillhörande slumpvariabler, samt händelser, för att svara på nedanstående frågor.

a) Du glömmer bort din telefon hemma, och kommer tillbaka till den efter två timmar. Hur många notifikationer kan du förvänta dig att du har fått under den här tiden? (2p)

b) Vad är sannolikheten att du fått minst 3 notifikationer? (1p)

c) När du återvänder visar det sig att du fått exakt 4 notifikationer. Vad är sannolikheten att alla notifikationerna är sms? (2p)

d) Du får ett sms kl 8:00. Vad är den förväntade tiden då du får nästa sms, varför? (1p)