**UPPSALA UNIVERSITET**
**Matematiska institutionen**
Patrik Tydesjö

Prov i matematik
SANNOLIKHET OCH STATISTIK
2021-06-02

*Skrivtid: 8 – 13 (plus 20 minuter för uppladdningsproceduren). Tillåtna hjälpmedel: Skrivdon, miniräknare, formelsamling samt tabellsamling. Samarbete mellan studenter eller mellan student och annan person är inte tillåtet. Lösningarna ska vara försedda med motiveringar. Varje uppgift ger högst 5 poäng. För betygen 3, 4 och 5 krävs totalt 18, 25 respektive 32 poäng. Ange din anonymitetskod på varje papper. På första bladet anges totalt antal sidor, och på varje blad anges sidnumrering. Saknar du anonymitetskod skriver du namn och personnummer på varje blad. Spara för säkerhets skull originalen till alla dina lösnngar. När du är klar skannar du in dina lösningar och skapar en pdf-fil. Allting ska vara i en och samma fil. Använd till exempel någon app. Ladda sedan upp filen på avsedd plats i Studium.*

**1.** Låt $a$ vara ett tal och låt $f$ vara en funktion sådan att $f(x) = a/x^2$ då $1 \leq x \leq 3$ samt $f(x) = 0$ för övrigt.

(a) Bestäm talet $a$ så att $f$ blir en täthetsfunktion. (2)

Låt nu $X$ vara den slumpvaraiabel som har $f$ som täthetsfunktion.

(b) Beräkna $P(X \leq 2)$. (1)

(c) Bestäm $E(X)$. (2)

**2.** Till Ångströmlaboratoriet kommer studenter med olika färdmedel. Andelen som promenerar är 2/10, andelen som cyklar är 5/10 och andelen som åker buss är 3/10. Sannolikheten att en student som promenerar till Ångströmlaboratoriet bor mer än 3 km från Ångströmlaboratoriet är 0.1, sannolikheten att en student som kommer på cykel bor mer än 3 km från Ångströmlaboratoriet är 0.4 och sannolikheten att en student som kommer med buss bor mer än 3 km från Ångströmlaboratoriet är 0.8.

(a) Beräkna sannolikheten att en slumpmässigt utvald student bor mer än 3 km från Ångströmlaboratoriet. (3)

(b) Man konstaterar att en viss utvald student bor mer än 3 km från Ångströmlaboratoriet. Vad är sannolikheten att denna student kommer med buss? (2)

**3.** En butik säljer en viss vara. Före varje arbetsdags början levereras 3 exemplar av varan till butiken. Antalet kunder som efterfrågar varan under en dag antas vara Poisson-fördelat med väntevärdet 2.

(a) Beräkna sannolikheten att varan under en viss arbetsdag räcker till för de kunder som efterfrågar varan. (2)

(b) Under en vanlig arbetsvecka (5 dagar), beräkna sannolikheten att varan räcker till minst 4 av dessa dagar. (3)

**4.** Ett åskväder passerar över Uppsala. Tiden mellan två blixtnedslag antas vara exponentialfördelad med väntevärdet 30 sekunder och dessa tider är oberoende av varandra. Om åskvädret pågår under 40 minuter, använd centrala gränsvärdessatsen för att approximativt beräkna sannolikheten att det blir minst 70 blixtnedslag. Ledning: Förutsättningen innebär att den totala tiden det tar för de 70 blixtarna att slå ned ska vara högst 40 minuter. (5)

**5.** Vi har ett slumpmässigt stickprov $x_1, x_2, x_3$ där motsvarande slumpvariabler $X_1, X_2, X_3$ är oberoende och har väntevärdet $\mu$ och variansen 2. Beteckna medelvärdet av observationerna med $\bar{x} = (x_1+x_2+x_3)/3$. På samma sätt har vi ett slumpmässigt stickprov $y_1, y_2$ där motsvarande slumpvariabler $Y_1, Y_2$ är oberoende och har väntevärdet $\mu$ och variansen 3. Medelvärdet av dessa observationer är $\bar{y} = (y_1 + y_2)/2$. Vidare är $Y_1, Y_2$ oberoende av $X_1, X_2, X_3$. Betrakta skattningarna

$$\hat{\mu}_1 = \frac{4\bar{x} + 5\bar{y}}{9}$$

och

$$\hat{\mu}_2 = \frac{2\bar{x} + 3\bar{y}}{5}.$$

(a) Visa att båda skattningarna är väntevärdesriktiga. (2)

(b) Vilken av skattningarna är mest effektiv? Motivera ditt svar. (3)

**6.** Man har ett slumpmässigt stickprov från $N(\mu, 2)$:

| | | | |
|---|---|---|---|
| 38.3 | 40.1 | 39.3 | 41.2 |

(a) Ta fram ett 95 % konfidensintervall för $\mu$. (3)

(b) Man upprepar proceduren 10 gånger, det vill säga att man gör nya stickprov i serier om 4 observationer och skapar 10 konfidensintervall utifrån dessa. Vad är sannolikheten att minst 9 av dessa intervall täcker det verkliga värdet på $\mu$? (2)

**7.** I en lågstadieklass vill man undersöka om en viss pedagogisk insats kan sägas ha någon effekt på elevernas förmågor. Man väljer ut 5 elever som får göra ett matematiktest. De gör ett test före den pedagogiska insatsen och ett (likvärdigt) test efter. Det är samma 5 elever vid de två tillfällena. Före insatsen får eleverna följande poäng på testet:

| Elev | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Poäng | 15 | 14 | 10 | 13 | 11 |

Efter insatsen får de följande poäng:

| Elev | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Poäng | 16 | 14 | 12 | 12 | 12 |

Man antar att poängfördelningen är normalfördelad.

(a) Skapa ett 95 % konfidensintervall för den förväntade skillnaden i poäng. (4)

(b) Kan man på den givna konfidensnivån säga något om huruvida den pedagogiska insatsen har någon inverkan på elevernas förmågor? Motivera ditt svar. (1)

**8.** Följande datamängd antas följa en linjär modell: $y_i = m + kx_i$.

| x | 2 | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|
| y | 0.8 | 2.2 | 2.5 | 4.5 | 4.7 |

(a) Beräkna skattningar av $k$ och $m$. (3)

(b) Beräkna förklaringsgraden $R^2$. Är det ett starkt linjärt samband? (2)