

Extra uppgifter SDA1

Mona Sfaxi

January 19, 2024

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

Uppgift 1

Låt X vara en variabel med n stycken observationer. Summan av X ges då av

$$\sum_{i=1}^n x_i \quad (1)$$

Lena befinner sig på ett bageri och ska köpa frallor till sig själv. Det finns fem olika sorter med olika priser.

Sort	havre	råg	vete	dinkel	blåbär
Styckpris	5	4	3	7	9

a.)

Använd formeln ovan i Ekvation 1 för att beräkna summan av kostnaden ifall Lena köper 1 av varje fralla.

b.)

Om k är en konstant så vet vi även att

$$\sum_{i=1}^n x_i \cdot k = k \sum_{i=1}^n x_i \quad (2)$$

Använd formeln i Ekvation 2 för att beräkna Lenas kostnad ifall hon köper 4 av varje fralla.

c.)

Summan av en kvadrerad variabel ges av

$$\sum_{i=1}^n x_i^2 \quad (3)$$

Lena ofta köpa frallor på bageriet under vintern. Men det brukar ofta vara väldigt kallt ute på vägen dit. Hon uppmätte temperaturen under 10 tillfällen och fick mätningarna.

-7 -3 1 1 -5 2 8 5 4 -2

Hon är intresserad av de kvadrerade avvikelserna från noll. Beräkna summan av de kvadrerade mätningarna med hjälp av Ekvation 3.

d.)

Istället för att titta på de kvadrerade avvikelserna från 0 än hon nu intresserad av att veta vad summan av de kvadrerade avvikelserna från medelvärde är, dvs

$$\sum_{i=1}^n (x_i - \bar{x})^2 \quad (4)$$

Där $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Beräkna denna summa åt henne.

Uppgift 2

Nedan ser vi en tabell över lite spelarstatistik för Leo Messi i Champions League säsongerna 15/16 till 22/23

Säsong	15/16	16/17	17/18	18/19	19/20	20/21	21/22	22/23
Spelade matcher	6	9	10	10	8	6	7	7
Mål	7	11	6	12	3	5	5	4
Assists	1	2	2	3	4	2	0	4
Minuter spelade	630	810	783	837	661	540	630	615

Källa: <https://www.transfermarkt.com/lionel-messi/leistungsdaten/spieler/28003/>

a.)

Beräkna det genomsnittliga antalet mål Messi gjort under dessa 8 säsonger.

b.)

Beräkna variansen och standardavvikelsen för det genomsnittliga antalet mål han gjort.

Uppgift 3

En florist får varje dag in en leverans med rosor. Han bestämmer sig en dag för att mäta 8 stycken av dem och får följande siffror i cm:

65 72 59 87 51 77 95 80

a.)

Räkna ut medelvärdet och medianen för rosorna.

b.)

Räkna ut standardavvikelsen och IQR.

c.)

Rita en boxplot över rosornas längd.

Uppgift 4

Figuren nedan visar en simultan frekvenstabell över två variabler, hur mycket man tränar per månad och ifall man brukar titta på skräckfilmer.

	Workout/month		
Scary movies	3 <	3-7	8 ≥
No	3	28	16
Yes	18	24	8

a.)

Hur stor andel tittar på skräckfilmer och tränar 3-7 gånger/ månad?

b.)

Ange marginalfördelningen för skräckfilmer och rita ett stapeldiagram utifrån detta.

c.)

Hur stor andel av dem som tränar minst 8 gånger per månad tittar på skräckfilmer?

Uppgift 5

Ett företag Kippy som tillverkar jordnötssmör vet att deras storsäljare: *jordnötssmörXXL* har en vikt som är normalfördelad med en medelvikt på 499g och en varians på 21.4.

a.)

Beräkna andelen som har en vikt mindre än 490g.

b.)

Beräkna andelen som har en vikt större än 502 g.

c.)

Beräkna andelen som har en vikt mellan 495 och 501g.

Uppgift 6

Kippy har blivit stämd för falsk marknadsföring då de hävdar att burkarna innehåller 500g jordnötssmör och många kunder inte tyckte att det stämde. En advokat tittar närmare på fallet och börjar räkna på några av siffrorna (där han utgår från modellen ovan). Han undrar bland annat:

a.)

Vad den högsta vikten måste vara på en burk för att den ska tillhöra de burkar med de 60% lägsta vikterna.

b.)

Vad vikten minst måste vara på en burk för att den ska tillhöra burkarna med de 12% högsta vikterna.

Uppgift 7

Nedan finns en tabell med två variabler *lön* och *månadssparande* som ska användas i en regressionsmodell där man vill titta på hur sparandet förhåller sig till lönen:

Lön	25010	31110	42329	21955	28712	35090	37214
Sparande	1050	2750	6400	2100	4800	5400	5100

a.)

Vilken variabel bör vara beroende variabel och vilken bör vara oberoende variabel? Motivera

b.)

Rita ett punktdiagram mellan den beroende och oberoende variabeln.

c.)

Räkna ut medelvärdet och standardavvikelsen för respektive variabel. Kan standardavvikelsen vara negativ?

d.)

Räkna även ut kovariansen (dvs $S_{x,y}$) mellan dem. Kan kovariansen vara större än 1? Kan den vara negativ?

e.)

Räkna ut korrelationen mellan de två variablerna. Kan korrelationen vara negativ? Kan korrelationen vara större än 1?

f.)

Räkna ut b_1 och b_0 och skriv ut regressionsekvationen. Tolka b_1 .

g.)

Vad blir R^2 -värdet? Ge en tolkning av värdet. Kan man räkna ut detta på ett annat sätt?

h.)

Gör 7 prediktioner för den beroende variabeln, utgå från värdena på den oberoende variabeln i tabellen ovan.

i.)

Vad blir residualvariansen? Vad blir residualstandardavvikelsen?

j.)

Vanligtvis bör detta aldrig göras men lek nu med tanken att du skulle göra tvärt emot det du gjorde ovan så att den oberoende variabeln du valde ovan blir den beroende istället. Om du då skattade en regressionsmodell, vad skulle då b_0 och b_1 bli? Vad skulle R^2 bli? Är resultatet konstigt?