

Statistiska Metoder

Laboration

Ali Leylani

2025-01-28

Syfte samt instruktioner

Målet med denna laboration är att säkerställa kunskap om teorin vi gått igenom fram till och med stickprovsfördelningar. Lösningarna till laborationen lämnas in i notebook-format genom ITHS-portalen. Vidare är detta en grupplaboration som med fördel löses tillsammans med ca 4-5 klasskamrater. Ni är således välkomna att diskutera frågorna fritt sinsemellan, men samtliga ska skriva och lämna in en egen notebook med lösningar.

Deadline för inlämning är onsdagen den 12:e februari, 23:59.

a)

Observera sannolikhetsfördelningen för den diskreta slumpvariabeln X nedan.

x	$P(X = x)$
1	0.05
2	0.10
3	0.15
4	0.05
5	0.05
6	0.05
7	0.05
8	0.05
9	0.40
10	0.05

Plotta denna sannolikhetsfördelning.

b)

I en separat graf, plotta den kumulativa fördelningssfunktionen $F(x)$.

c)

Beräkna $E(X)$, $V(X)$ samt $S(X)$.

d)

Din uppgift är nu att ta stickprov på X , och beräkna stickprovsmedelvärden. Syftet är att att se om vi kan hitta några mönster. Vi kommer börja med små stickprovsstorlekar och efterhand öka storleken.

1. Ta tusen stickprov med $n = 1$, från fördelningen X .

tips: `random.choices` kan vara smart att använda här.

2. Registerara medelvärdet för samtliga av dina stickprov, enskilt.

obs: för $n = 1$ är det inte nödvändigt att ta medelvärdet (**varför?**)

3. Plotta alla dina medelvärden med ett histogram.

4. Repetera 1-3 men med $n = 2, 4, 8, 12, 15, 20, 24, 30, 35, 50, 100, 200$ istället. Du bör nu ha totalt 13 histogram med fördelningar för dina stickprovsmedelvärden.

tips: Det kan vara en god idé att justera x lim för graferna, för att få jämförbara plottar.

e)

Låt oss nu försöka förstå oss på resultaten från d).

Hur tolkar du plotten för $n = 1$? Vad är det som händer då n växer? Vilken fördelning börjar dina histogram att likna?

obs: dina y-axlar visar nu absolut frekvens, snarare än relativ frekvens. Det går att med ett argument till `plt.hist` istället plotta relativ frekvens.

f*)

För vilket värde på n är det lämpligt att använda Z -fördelningen för analyser av stickprovsmedelvärdet \bar{X} ? För vilket värde på n är det lämpligt att använda t -fördelningen för ditto?