

4.2 - Confidence intervals - Confidence intervals for mean

Q1 The following random sample is taken from a Normal distribution with standard deviation 0.22 .

0.958, 1.219, 1.018, 0.999, 1.531, 1.338, 1.501

a) What is the sample mean?

SVAR:

Summera ihop alla termer och dela på antalet termer

b What is the standard deviation of the sample mean of random samples of 7?

SVAR:

$\sqrt{0.22^2 / (\text{antal } x)}$

c) Based on the sample, what is the lower 95% confidence limit for the population mean?

SVAR:

$\text{mean}(x) - (Z * \text{sDev} / \sqrt{\text{length}(x)})$

Z:

80%	1.28
85%	1.44
90%	1.64
95%	1.96
98%	2.33
99%	2.58

d) And what is the upper 99% confidence limit for the population mean?

SVAR:

$$\text{mean}(x) + (Z * sDev / \sqrt{\text{length}(x)})$$

Q4 The sample estimate s of standard deviation (σ) is reliable when the sample is big enough. (What is 'enough' depends on the population; but 100 and over is usually enough). In such cases we often use s for σ in calculating a confidence interval. A sample of 200 gives a mean of 1.91 and variance $s^2=0.4$. Calculate the 89 % confidence interval for the population mean (μ).

Intervallens nedre gräns:

SVAR:

$$CI = \text{mean}(x) - Z * \sqrt{\text{Var}}/\sqrt{n}$$

Intervallens övre gräns:

SVAR:

$$CI = \text{mean}(x) + Z * \sqrt{\text{Var}}/\sqrt{n}$$

Q5 Följande stickprov är taget från en normalfördelning med standardavvikelse 7.8 .
Bestäm en 90% konfidensintervall för populationsmedelvärdet.

23.1, 33.5, 17.7, 37.1, 25.4, 36.1, 19.4, 30.6, 10.3

Intervallens nedre gräns:

SVAR:

$$CI = \text{mean}(x) - Z * sdev/\sqrt{n}$$

Intervallens högre gräns:

SVAR:

$$CI = \text{mean}(x) + Z * sdev/\sqrt{n}$$

Q6 Antalet timmar som studenter deltidsarbetar är normalfördelat med standardavvikelse 9 timmar. Ett stickprov på 150 studenter togs och stickprovets medelvärde beräknades till 12.3 timmar.

SVAR:

Se lösning på tidigare uppgifter.

Q7 En biverkning av att sluta röka är att man går upp i vikt. Viktökningen antas vara normalfördelad med standardavvikelse 2 under det första året. Ett stickprov på 38 personer som nyligen slutat röka deltar i studien, och deras genomsnittliga viktökning under året är 3.5 kg.

SVAR:

Se lösning på tidigare uppgifter.

Q8 Vi har ett stickprov på 18 individer då man testat ett träningsredskap för uthållighet på ett gym. Antalet minuter tills övervakningsinsapparaten avbryter övningen blev i snitt för deltagarna 15 minuter. Standardavvikelsen för genomsnittliga tiden beräknades vara: 2.8 minuter. Under förutsättningen att tiden för övningen är normalfördelad, beräkna en 90 % konfidensintervall för populationens medeluthållighet i minuter enligt denna övning.

a) Nedre gräns för konfidensintervallet

SVAR:

$$CI = \text{mean}(x) - t * \text{sdev}/\sqrt{n}$$

Talet t hittar man från menyn: Help Tables Inverse T (Both) med $p = \alpha$ och $\nu = n - 1$.

b) Övre gräns för konfidensintervallet:

SVAR:

$$CI = \text{mean}(x) + t * \text{sdev}/\sqrt{n}$$

Q9 Följande siffror visar temperaturen i grader celsius av reaktorkärnan för en kemisk reaktor, tagna vid olika tillfällen under en månads tid. Temperaturen anses vara normalfördelad.

Katalysatorns effektivitet försämrar snabbt om temperaturen i kärnan överstiger 845 grader celsius. Hitta en 99% konfidensintervall för medeltemperaturen av reaktorkärnan.

735

779

798

864

844

900

813

792

846
796
832
826
801
919
837
821
799

SVAR:

för att få T: table -> inverse T -> both:

p = 1-kondensintervallsgraden(procent).

V = n-1

$x \pm t * sdev / \sqrt{n}$

Q10 Följande data kommer från en undersökning bland brittiska studenter om deras deltidsarbeten. Data visar den genomsnittliga veckolönen för respektive student med deltidsarbete, i engelska pund. Bestäm en 95% konfidensintervall för medelvärdet av studenternas veckolön.

52
62
51
70
43
71
59
71
43
64
59
58
67
40
69
56
57
43
43
68

SVAR:

Se lösning på Q9

Q11 En viss population följer normalfördelningen med variansen 327 och ett okänt medelvärde. Man vill ta ett stickprov för att få en approximation till medelvärdet. Hur stort stickprov behöver man, så att approximationen inte hamnar längre än 2 från det riktiga medelvärdet:

SVAR:

$$(Z * \sqrt{\text{Var}}) / B)^2$$

Z hämtas från VLE-table.

B = hur stort felet kan bli. 2 i det här fallet.

OBS: Avrunda till närmast större heltal om svaret är decimalt.

Q12 En studie genomfördes där man ville skatta väntevärdet för en normalfördelad variabel med standardavvikelsen 4.9. Efter att ha undersökt 90 enheter så erhöll man konfidensintervallet:

$$91.6 \leq \mu \leq 93.4$$

a Vilken konfidensgrad har intervallet?

Ange svaret i heltalsprocent.

SVAR:

$$m = (93.4 - 91.6) / 2$$

$$Z = (\sqrt{n} * m) / \text{sdev}$$

table - normal both - x = z

grad = 1 - X (X = resultatet ur table. Notera: inte lilla x, utan stora)

b Hur många observationer borde man ha undersökt om man vill ha samma intervallängd men en 90%-ig konfidensgrad?

Avrunda uppåt till närmaste heltal!

SVAR:

$$m = (93.4 - 91.6) / 2$$

$$\text{wolphramalpha} : Z * (\text{sdev} / \sqrt{n}) = m$$

4.4 - Central Limit Theorem

Q1 Previous research has estimated that 20% of passengers travelling from the USA to Sweden carry more than their duty-free allowance. A plane carrying 90 passengers arrives at Landvetter airport and customs officers randomly select 18 passengers and search them.

SVAR:

a) b) c)

I MatLab:

format long

binopdf(6, 18, .2)

Q2 Innehållet i en säck 'ballast' (bergmaterial, krossad sten) för betongtillverkning är normalfördelat med medelvikten 25 kg och en standardavvikelse på 1.3 kg.

a Vad är sannolikheten för en slumpmässigt vald säck att dess vikt understiger 24.6 kg ?

SVAR:

mean = 25

sdev = 1.3

table low, $P(X < 24.6)$

b Vad blir sannolikheten för 8 slumpmässigt valda säckar att säckmedelvikten understiger 24.6 kg.

SVAR:

sdev = $1.3/\sqrt{8}$

som a)

c Vad är sannolikheten för en slumpmässigt vald säck att dess vikt ligger mellan 24.6 och 25.6 kg ?

SVAR:

mean = 25

sdev = 1.3

table low, $P(X < 25.6) - P(X < 24.6)$

d Vad blir sannolikheten för 22 slumpmässigt valda säckar att säckmedelvikten ligger mellan 24.6 och 25.6 kg?

SVAR:

mean = 25

sdev = $1.3/\sqrt{22}$

table low, $P(X < 25.6) - P(X < 24.6)$

Q3 Beloppet som en slumpmässigt vald kund spenderar per besök på ett nöjesfält är i genomsnitt 301 kronor med en standardavvikelse på 87 kronor.

Hur stor är sannolikheten att en slumpmässigt vald grupp på 95 individer spenderar fler än 28400 kronor på ett besök?

SVAR:

mean = $301 * 95$

sdev = $\sqrt{95} * 87$

table high, $P(X > 28400)$

Q4 Antag att diametern hos tillverkade muttrar har en standardavvikelse på 0.45 mm. För utvärdering av tillverkningsprocessen genomförs en kvalitetskontroll varvid man tar ut 3 muttrar och mäter deras diameter. .

Hur stor är risken att den genomsnittliga diametern i urvalet avviker från diameterns väntevärde med mer än 0.1125 mm?

SVAR:

De tillverkade skruvarnas diameter d har en okänd fördelning med okänt väntevärde μ standardavvikelse $\sigma = 0.45$.

Enligt uppgiften vet vi att $\bar{d} - \mu = 0.1125$ då vi har ett stickprov på 35 skruvar. Centrala gränsvärdessatsen ger oss att:

$$Z = \frac{\bar{d} - \mu}{\frac{\sigma}{\sqrt{35}}} \sim N(0, 1) \text{ fördelad.}$$

Vi söker nu sannolikheten

$$P(Z < \frac{-0.1125}{\frac{0.45}{\sqrt{35}}}) + P(Z > \frac{0.1125}{\frac{0.45}{\sqrt{35}}})$$

5.1 - Proportions - Approximations

Q1 The probability of success on any trial of a binomial experiment is 45 % .

Use the Normal approximation to the binomial to find the probability that the proportion of successes in a sample of 800 is less than 42 %.

(Do not make use of a "continuity correction" in this question.)

Svar:

vle table -> normal low

p = sannolikheten = 0.45

n = antal av nagot = 800

y = sannolikheten mindre än = 0.42

$$\begin{aligned} \text{mean} &= n * p &= 0.45 * 800 &= 360 \\ \text{sdev} &= \sqrt{n * p(1 - p)} &= \sqrt{800 * 0.45 * (1 - 0.45)} &= 14.0724728 \\ x &= y * n &= 800 * 0.42 &= 336 \end{aligned}$$

$$p(X < 336) = 0.04405$$

Q3 In the last election 61 % of the electorate voted for party P.

- a On the assumption that the electorate have not changed their allegiances what is the probability that in a random sample of 50 electors, 27 (54%) or fewer indicate that they will vote for party P. (Use the binomial distribution).

SVAR:

I MatLab:


```
x = 0:27
sum(binopdf(x; 50; 0.68))
svar: 0.191551320999700
```

b Use the Normal approximation to the binomial distribution to estimate the probability that 27 people or fewer in a random sample of 50 indicate that they will vote for party P. Use a "continuity correction".

SVAR:

ingen binomialsak när man har continuity-värden...

mean = $n * p = (50 * 0.61)$

sdev = $\sqrt{n * p * (1 - p)} = \sqrt{50 * 0.61 * (1 - 0.61)}$

$x = 27 \pm$ correction enligt tabell (i detta fallet + 0.5).

IF	use
$P(X=n)$	$P(n - 0.5 < X < n + 0.5)$
$P(X>n)$	$P(X > n + 0.5)$
$P(X \leq n)$	$P(X < n + 0.5)$
$P(X < n)$	$P(X < n - 0.5)$
$P(X \geq n)$	$P(X > n - 0.5)$

källa: <http://www.statisticshowto.com/articles/what-is-the-continuity-correctionfactor/>

Table -> Normal Low -> sätt in ovanstående värden.

svaret i decimalform: 0.19219

c In a large sample the Normal distribution is a good representation of the probability distribution of the proportion (supporting party P). If the sample size were 400 electors, we would use the Normal distribution with mean 0.61 and what standard deviation?

SVAR:

sdev = $\sqrt{n * p * (1 - p)} / n = \sqrt{400 * 0.61 * (1 - 0.61)} / 400 = 0.0243874967965$

d On this basis find the probability that 54% or less of the large sample intend to vote for party P. Do *NOT* make use of a "continuity correction".

SVAR:

vle table -> normal low:

$x = .54$

mean = $p * n = .54 * 400 = 216$

sdev = $\sqrt{n * p * (1 - p)} = \sqrt{400 * 0.54 * (1 - 0.54)} = 9.9679486355$

returneraX

Q4 Given a binomial random variable X with $n=31$ and $p=0.36$, find the (exact) probabilities of the following events.

a $P(X = 12)$

SVAR:

$P(X = k)$.

MATLAB: `binopdf(k, n, p)`

b $P(X \leq 5)$

SVAR:

$P(X \leq k)$.

`binocdf(k, n, p)`

c Find the Normal approximation to

$P(X=12)$. Use the "continuity correction".

SVAR:

$P(X=k)$. Use the "continuity correction" .

use table for "continuity correction"

table \rightarrow normal law :

mean = $n * p = 11.16$

sdev = $\sqrt{n * p * (1-p)} = 2.67252689416$

$P(X < 12.5) = 0.69195$

$P(X < 11.5) = 0.55062$

$0.69195 - 0.55062 = \mathbf{0.14133}$

d Find the Normal approximation to

$P(X \leq 5)$. Use the "continuity correction".

SVAR:

table \rightarrow normal law

$P(X < 5.5) = 0.01709$

e Is the answer to (c) about the same as the answer to (a)? Is the answer to (d) about the same as the answer to (b)?

SVAR:

(c) near (a), and (d) near (b)

Q5 I en fabrik tillverkar man komplicerade elektronikkomponenter. Andelen defekta komponenter är 3.5% av totala produktionen. Ett stickprov på 800 komponenter väljs slumpmässigt ut till test. Förutsätt att andelen defekta komponenter är samma i stickprovet som i totala produktionen nämligen: 3.5%.

a Vad blir sannolikheten att 35 eller fler av komponenterna i stickprovet är defekta? Använd normalapproximation **utan halvkorrektion** i denna del!

SVAR:

t a b l e -> normal high :

$$\text{mean} = n * p = 800 * 0.035 = 26$$

$$\text{sdev} = \sqrt{n * p * (1 - p)} = 5.19807656735$$

$$x = 35$$

$$P(X \geq 35) = 0.08905$$

b Räkna på samma sannolikhet som i uppgift a) och använd normalapproximation **med halvkorrektion**.

Samma som förra men

$$P(X \geq \mathbf{34.5}) = 0.07453$$

c Räkna på samma sannolikhet som i uppgift a) och använd poissonapproximation. Använd excels funktion för poissonfördelning.

SVAR:

<http://faculty.vassar.edu/lowry/poisson.html> // Tips: Använd nästa länk ifall du vill ha normal/poisson/exakt binomial

$$n = 800$$

$$k = 35$$

$$p = 0.035$$

klicka på calculate och kolla vilket svar som passar din fråga. I detta fall så söker vi p för 35 eller fler, så svaret blir : 0.112101

d Slutligen räkna på samma sannolikhet som i uppgift a) exakt. Använd excels funktion för binominalfördelning.

SVAR:

<http://www.vassarstats.net/binomialX.html> - JÄVLIGT BRA

$$n = 800$$

$$k = 35$$

$$p = 0.035$$

P: 35 or more out of 800

Method 1. exact binomial calculation: 0.108106398184

Q6 I en serie försök är sannolikheten för lyckad resultat 23 % för de individuella experimenten.

Använd normalapproximationen av binominalfördelningen för att hitta sannolikheten att mindre än 22 % av experimenten lyckats då man utför totalt 800 försök.

(Använd **inte** halvkorrektion i denna uppgift.)

SVAR:

$$\text{Väntevärde/mean} = 0.23$$

$$\text{sdev} = \sqrt{(0.23)(1-0.23)/800}$$

Man räknar ut variansen genom att ta $(\text{mean})(1-\text{mean})/n$ och får standardavvikelsen genom att ta roten ur på variansen.

table -> low normal
 $P(X < 0.22) = 0.25076$

Q7 Låt X vara en binominalfördelad stokastisk variabel med $n=34$ och $p=0.45$.
Beräkna (exakt) följande två sannolikheter:

a $P(X = 20)$.

SVAR:

<http://www.vassarstats.net/binomialX.html>

$n = 34$

$k = 20$

$p = 0.45$

P: exactly 20 out of 34 - Method 1

b $P(X \leq 14)$

SVAR:

<http://www.vassarstats.net/binomialX.html>

$n = 34$

$k = 14$

$p = 0.45$

P: 14 or fewer out of 34 - Method 1

c Hitta Normalapproximationen för sannolikheten:

$P(X=20)$. Använd halvkorrektion.

SVAR:

mean = $n * p = (34 * 0.45) = 15.3$

sdev = $\sqrt{n * p * (1 - p)} = \sqrt{(0.45 * 34)(1 - 0.45)} = 2.90086194087$

$P(X < 20.5) = 0.96348$

$P(X < 19.5) = 0.92617$

$0.96348 - 0.92617 = 0.03731$

d Hitta normalapproximationen för sannolikheten:

$P(X \leq 14)$. Använd halvkorrektion.

SVAR:

IF use
P(X=n) P(n - 0.5 < X < n + 0.5) == P(X < n-0.5) - P(X < n+0.5)
P(X>n) P(X > n + 0.5)
P(X <= n) P(X < n + 0.5)
P (X<n) P(X < n - 0.5)
P(X >=n) P(X > n - 0.5)

källa: <http://www.statisticshowto.com/articles/what-is-the-continuity-correctionfactor/>

$$\text{mean} = n * p = (34 * 0.45) = 15.3$$

$$\text{sdev} = \sqrt{n * p * (1 - p)} = \sqrt{(0.45 * 34)(1 - 0.45)} = 2.90086194087$$

$$P(X \leq n) \rightarrow P(X < n + 0.5)$$

table -> normal low

$$P(X < 14.5) = 0.39136$$

5.1 - Proportions - Confidence intervals

Q1 Vid en opinionsundersökning uppskattas stödet för ett politisk parti. I ett stickprov på 676 människor har 191 uttryckt sitt stöd till fördel för partiet. Beräkna ett 95% och 99% confidensintervall för andelen (uttryckt i siffror mellan 0 och 1) som stödjer partiet.

SVAR:

$$CI = p \pm Z * \sqrt{p(1 - p) / n}$$

$$p = 191 / 676$$

$$n = \text{totala} = 676$$

Z:

80%	1.28
85%	1.44
90%	1.64
95%	1.96
98%	2.33
99%	2.58

a Nedre intervallgräns för 95% .

$$0.2486 = 0.2825 - 1.96 * \sqrt{(0.2825 * (1 - 0.2825))/676}$$

b Övre intervallgräns för 95% .

$$0.3164 = 0.2825 + 1.96 * \sqrt{(0.2825 * (1 - 0.2825))/676)}$$

c Nedre intervallgräns för 99% . **d** Övre intervallgräns för 99% .

Gör som ovan men byt ut Z värdet.

Q2 I en undersökning med slumpmässigt valda studenter, tog man ett stickprov på 250 elever. Varje student frågades om hon/han skulle nu ha valt en annan kurs. I resultatet antecknades en 1:a om svaret var ja och 0:a om svaret var nej. Beräkna en 90% konfidensintervall för andelen elever som har ångrat sina val av kurs.

SVAR:

Samma princip som förra uppgiften.

a Nedre gräns av intervallet.

$$(70/250) - 1.64 * \sqrt{((70/250) * (1 - (70/250)))/250} = 0.2334$$

b Övre gräns av intervallet.

$$(70/250) + 1.64 * \sqrt{((70/250) * (1 - (70/250)))/250} = 0.3265$$