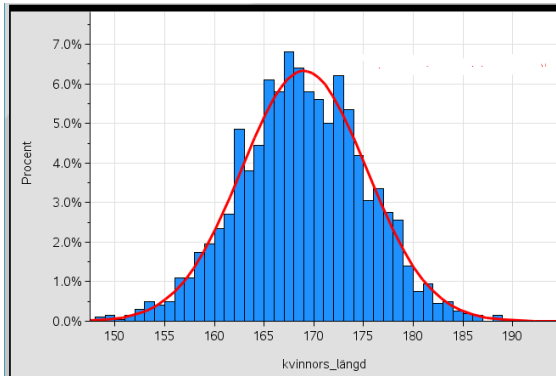


## Arbeta med normalfördelningar

I en större undersökning om hur kvinnors längd gjorde man undersökning hos kvinnor i ett viss åldersintervall. Man drog sedan ett slumpmässigt urval på 2000 kvinnor och resultatet blev ett medelvärde på 169,0 cm och en standardavvikelse på 6,3 cm. Fördelningen syns i diagrammet nedan.

Man kan se att fördelningen ganska väl följer en s.k. klockkurva.

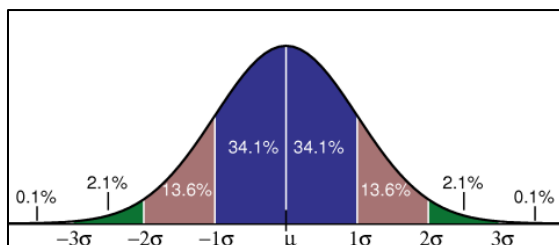


Statistiska data som ger en sådan här fördelning kallas *normalfördelning*. Många företeelser i naturen och i samhället följer approximativt en normalfördelning.

Längden hos kvinnorna har en ganska stor spridning. De allra flesta är placerade i mitten nära medelvärdet men det finns både långa och korta.

För *alla* normalfördelning gäller att ca 68 % av kvinnorna har en längd inom en standardavvikelse från medelvärdet, dvs. i intervallet  $169,0 \pm 6,3$  cm (dvs. mellan 162,7 och 175,3 cm).

Diagrammet nedan visar hur stor andel som ligger 1, 2 och 3 standardavvikelser från medelvärdet i en teoretisk normalfördelning.



### Analys av data från Högskoleprovet

Vi ska nu titta närmare på data som vi har laddat ner till räknaren med programmet *TI Connect*. Dessa data bifogas denna aktivitet som två listor.

Det gäller resultatet från Högskoleprovet våren 2015. Antalet deltagare var 75 471 och medelvärdet i det normerade provet var **0,90** och standardavvikelsen **0,39**.

Data har vi hämtat från <http://studera.nu/hogskoleprov/fpn/normeringstabeller-varen-2015/>

Maxpoäng på provet är 2,0 och resultaten gavs i intervall med 0,05 poäng. Titta på data i statistikeditorn nedan.

NORMAL FLYT AUTO REELL RAD MP					
POANG	ANTAL	L1	L2	L3	3
0	269				
0.05	397				
0.1	578				
0.15	827				
0.2	1055				
0.25	1166				
0.3	1386				
0.35	1650				
0.4	1860				
0.45	2060				
0.5	2348				

L1(1)=

Vi ska nu först räkna om våra absoluta värden (ANTAL) till relativa värden, dvs. andelar. Placera då markören högst upp i kolumnhuvudet i lista L1 och tryck på **enter**. Skriv sedan in enligt nedan.

NORMAL FLYT AUTO REELL RAD MP					
POANG	ANTAL	L1	L2	L3	3
0	269				
0.05	397				
0.1	578				
0.15	827				
0.2	1055				
0.25	1166				
0.3	1386				
0.35	1650				
0.4	1860				
0.45	2060				
0.5	2348				

L1= LANTAL/75471

Listan poäng kopierar du in genom att trycka på **2nd**[list] och sedan välja POANG i namnlistan. När du är klar trycker du på **stat** igen.

Se nästa sida.

POANG	ANTAL	L1	L2	L3	3
0	269	0.0036	-----	-----	
0.05	397	0.0053			
0.1	578	0.0077			
0.15	827	0.011			
0.2	1055	0.014			
0.25	1166	0.0154			
0.3	1386	0.0184			
0.35	1650	0.0219			
0.4	1860	0.0246			
0.45	2060	0.0273			
0.5	2348	0.0311			

L1(1)=0.00356428296961

L1 visar nu andelarna. 0.0036 betyder då 0,36 % osv.

Vi kan nu rita ett diagram över fördelningen. Vi trycker på  $\text{2nd}$ [stat plot] för att göra diagraminställningen.

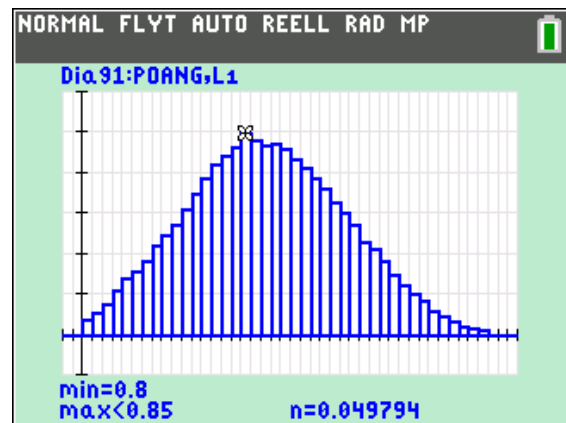
NORMAL FLYT AUTO REELL RAD MP	
TRYCK [◀]ELLER[▶] FÖR VAL AV ALTERNATIV	
Diq.91	Diq.92 Diq.93
Pa Av	
Skriv:	
Xlista	: POANG
Frekv	: L1
Färg	: BLÅ

Vi ska alltså rita ett histogram och vi har POANG på x-axeln och frekvensen för poängen på y-axeln.

Inställning av ett bra fönster:

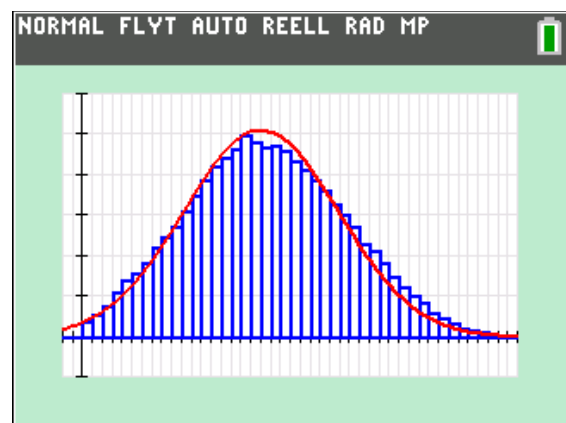
NORMAL FLYT AUTO REELL RAD MP	
FÖNSTER	
Xmin	= -0.1
Xmax	= 2.2
Xsk1	= -0.05
Ymin	= -0.01
Ymax	= 0.06
Ysk1	= 0.01
Xres	= 1
$\Delta X$	= 0.00871212121212
SpåraSteg	= 0.017424242424...

Vi ser att fördelningen påminner om en normalfördelning.



Fördelning med andel på y-axeln

Vi lägger nu på en normalfördelningskurva i samma diagram. Den överlappar våra data utomordentligt bra.



Det här betyder att vi i fortsättningen kan använda normalfördelningskurvan i stället för data i fortsatta beräkningar.

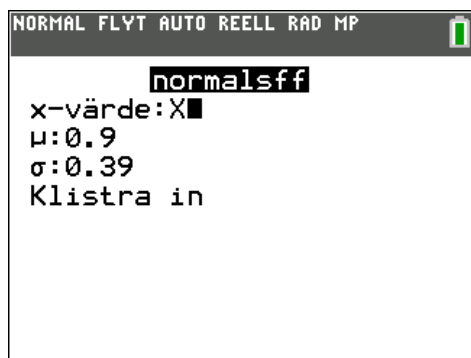
Räknaren har ett antal verktyg för att göra beräkningar på normalfördelningar. Tryck nu först på  $\text{2nd}$ [distr]. Då öppnas en meny. Under FÖRDELN finns beräkningsverktyg för normalfördelningar. Det är de tre första.

NORMAL FLYT AUTO REELL RAD MP	
FÖRDELN RITA	
1:	normalsff(
2:	normalkff(
3:	invNorm(
4:	invT(
5:	tsff(
6:	tkff(
7:	$\chi^2$ sff(
8:	$\chi^2$ kff(
9↓	Fsff(

Vi går igenom dem, en i taget.

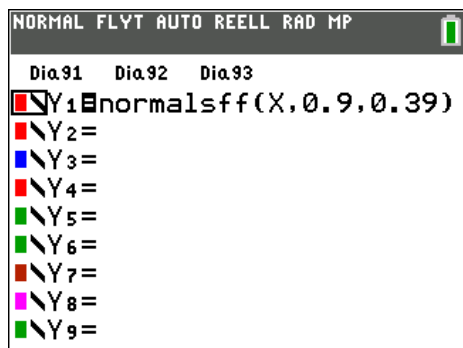
**1:normalsff** (förkortning för normala frekvensfunktionen)

1. Placera markören i grafitningseditorn. Tryck på  $\boxed{\vee}$  och placera markören på första raden
2. Tryck nu på  $\boxed{2nd}$  $\boxed{[distr]}$  och välj alternativ 1. Då kommer en meny upp där du fyller i enligt nedan.

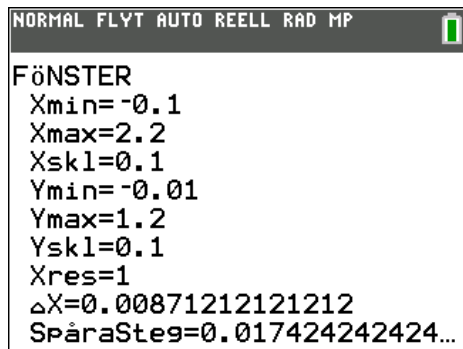


Vi använder detta för att rita en normalfördelningskurva. Variabeln är X och sedan fyller man i medelvärde och standardavvikelse.

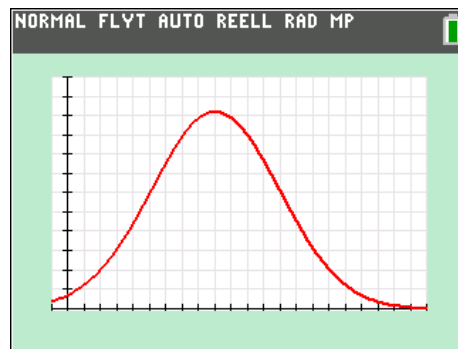
3. Gå ner till Klistra in och tryck på  $\boxed{enter}$ . Nu skrivs ett uttryck in på inmatningsraden i grafitningseditorn.



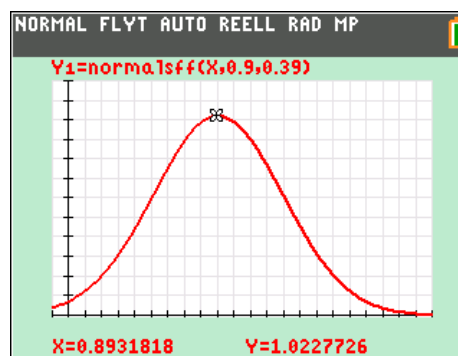
Dags att rita. Vi måste nu välja ett bra fönster för att kunna visa grafen i dess helhet. Här är en bra inställning:



Dags att plotta kurvan. Tryck på  $\boxed{graph}$ .



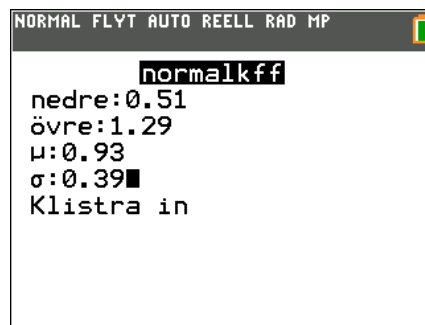
Om vi spårar i kurvan (tryck på  $\boxed{trace}$ ) ser vi att det största värdet antas när x är ca 0,90.



Vi ser också en annan sak: kurvan planar inte ut mot 0 utan antar också värden när  $x < 0$ . Detta gäller alla normalfördelningar: de har *oändlig* utsträckning i x-led.

**2:Normalkff** (förkortning för normala kumulativa frekvensfunktionen)

1. Placera markören i grundfönstret.
2. Tryck nu på  $\boxed{2nd}$  $\boxed{[distr]}$  och välj alternativ 2. Då kommer en meny upp där du fyller i enligt nedan.



Man fyller i nedre och övre gräns för beräkningarna. I detta fall vill vi beräkna andelen som har ett resultat som ligger högst *en standardavvikelse* ifrån medelvärdet ( $0,90-0,39=0,51$ ,  $0,90+0,39=1,29$ ).

Markera Klistra in och tryck på  $\boxed{enter}$ .

```
NORMAL FLYT AUTO REELL RAD MP
normalkff(0.51,1.29,0.93,▶
.....
0.6812591155
```

Nu får vi direkt andelen med resultat inom en standardavvikelse från medelvärdet, 0,68 eller 68 %. Jämför med figuren på sid 1 där vi ser att det gröna området är just 68 %.

Om vi vill beräkna hur stor andel som har ett resultat mellan 0,75 och 1,25 skriver vi så här:

```
NORMAL FLYT AUTO REELL RAD MP
normalkff
nedre:0.75
övre:1.25
μ:0.93
σ:0.39
Klistra in
```

Vi markerar Klistra in och trycker på `enter`.

Uttrycket klistras in i grundfönstret och vi trycker återigen på `trace`. Vi får svaret 0,47 eller 47 %.

```
NORMAL FLYT AUTO REELL RAD MP
normalkff(0.75,1.25,0.93,▶
.....
0.4718319669
```

Vi kan alltså beräkna andelen i ett valfritt intervall.

En snabbkoll mot våra data i statistikeditorn:

POANG	ANTAL	L1	L2	L3	3
0.75	3509	0.0465			
0.8	3758	0.0498			
0.85	3619	0.048			
0.9	3537	0.0469			
0.95	3570	0.0473			
1	3473	0.046			
1.05	3283	0.0435			
1.1	3126	0.0414			
1.15	2926	0.0388			
1.2	2731	0.0362			
1.25	2461	0.0326			

L1(16)= 0.04649468007579

Om man summerar alla data som syns i kolumn L1 i fönstret ovan så får man också ca 0,47.

### 3:InvNorm (förkortning av *inversa* eller omvända normalfördelningen)

Detta verktyg används när man vill beräkna den *övre* gränsen för att få en viss andel. Antag att vi vill beräkna den övre gränsen så att man täcker in 90 % av alla deltagare.

Tryck på **3:Invnorm** i grundfönstret och skriv sedan in enligt nedan:

```
NORMAL FLYT AUTO REELL RAD MP
invNorm
area:0.9
μ:0.90
σ:0.39
Klistra in
```

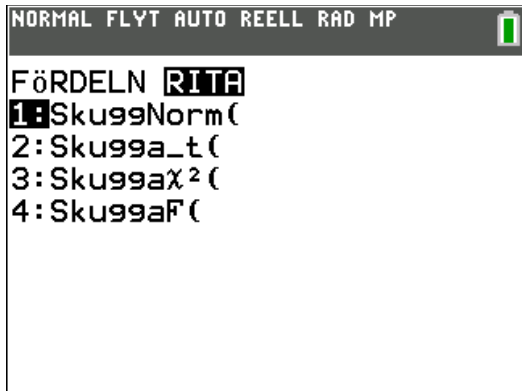
Markera Klistra in och tryck på `enter`. Uttrycket kopieras in i grundfönstret. Tryck nu på `enter` igen:

```
NORMAL FLYT AUTO REELL RAD MP
invNorm(0.9,0.90,0.39)
.....
1.399805111
```

Vi får ca 1.40 som resultat. Det betyder att ca 90 % av deltagarna har maximalt 1,40 poäng.

Vi räknar åt andra hållet här: Vi känner en andel och ska beräkna en gräns i ett intervall.

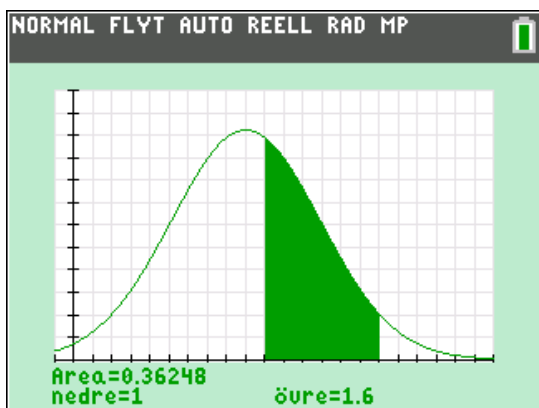
Till sista ska vi titta på ett mer grafiskt verktyg. Tryck på  $2^{nd}$ [distr] igen och välj RITA och sedan 1:SkuggaNorm.



Då fyller du in enligt nedan.



Vi ska nu beräkna andelen som har ett resultat mellan 1 och 1,6. Markera Rita och tryck  $\text{enter}$ . Nu plottas normalfördelningskurvan upp och intervallet markeras i färg. Du ser resultatet längst ner. Vi ser att 36 % har ett resultat mellan 1 och 1,6.



Om du vill göra en ny beräkning med andra värden *måste* du trycka på  $2^{nd}$ [draw] för att radera den nuvarande plottningen.

## Analys av vikten hos krossade tomater

Vi har här tagit med en uppgift från ett nationellt prov och vi ska nu använda räknarens verktyg för normalfördelning för att lösa uppgifterna.

**20.** Ett företag fyller konservburkar med krossade tomater. Enligt märkningen innehåller en burk 400 g tomater. Tomaternas vikt är normalfördelad kring medelvärdet 395 g och standardavvikelsen är 5,0 g.

a) Hur många procent av konservburkarna kan förväntas innehålla mindre än de 400 g som anges på burken? (2/0/0)

Företaget vill inte ha för många missnöjda kunder och tänker därför fylla konservburkarna lite mer. De ändrar kravet till att minst 97,7 % av burkarna ska innehålla minst 400 g tomater. Standardavvikelsen antas fortfarande vara 5,0 g.

b) Beräkna vilket medelvärde på vikten som motsvarar detta nya krav. (0/3/0)

a) Man kan ju börja med att plotta fördelningen. Vi känner ju till medelvärde och standardavvikelse.

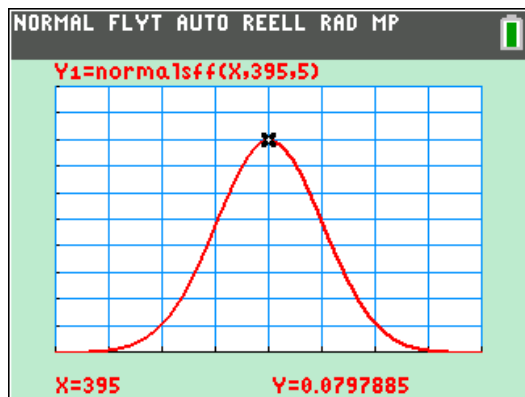
Placera markören i Y1 i inmatningsrutan för funktioner. Tryck på  $2^{nd}$ [distr] och välj 1:normalsff. Fyll sedan i enligt nedan.



Klistra nu in detta i Y1. Nu kan vi plotta kurvan. Vi måste naturligtvis se till att vi har ett bra fönster. Eftersom medelvärdet är 395 ska det ligga någonstans i mitten och då är ett lämpligt intervall för X 375 - 415. Då ligger medelvärdet precis i mitten.

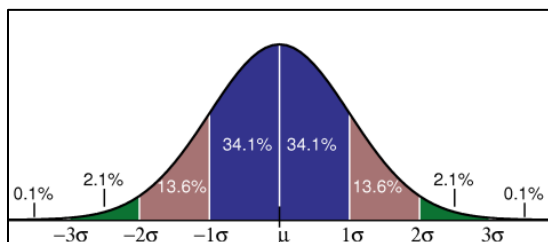
Att ställa in intervallet i y-led är något knepigare och man kan få pröva sig fram. Om man sätter Ymax till 0,1 blir det bra.

Nu plottar vi!



När du kommit så här långt känner du säkert till de värden som gäller för *alla* normalfördelningar. Dessa värden finns ju också med i den formelsamling som man får använda på prov.

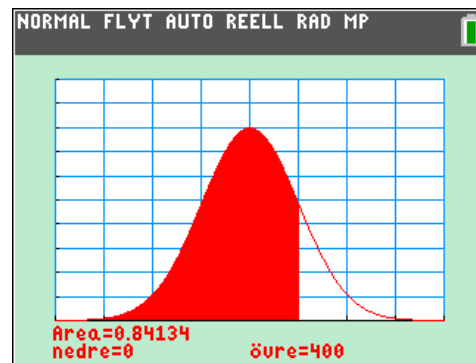
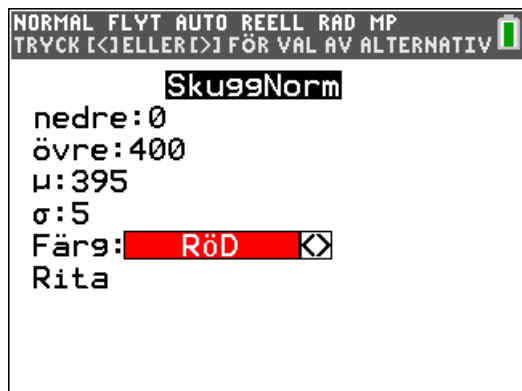
På sid 1 hade vi följande figur:



Du ska alltså beräkna andelarna från vänster och fram till 400.

Det är 50 % + 34,1 % = 84,1 %.

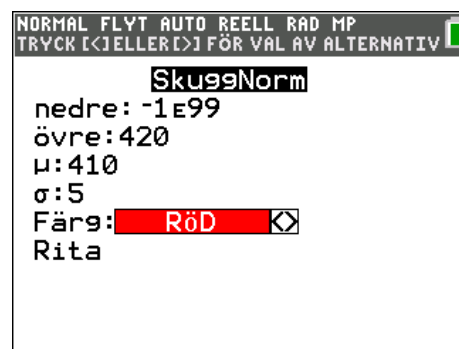
Om du är osäker kan du kontrollera med räknaren. Vi har här använt RITA-funktionen som vi gick igenom.



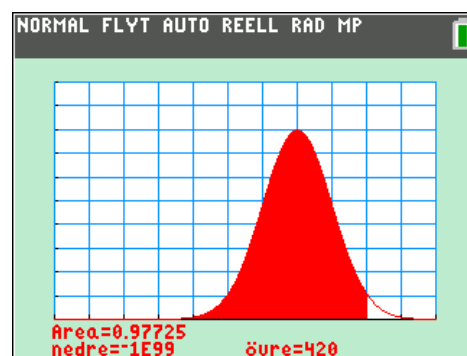
b) Den här uppgiften är något knepigare. 97,7 % av burkarna ska innehålla minst 400 g tomater. Vi tittar på figuren med värdena igen. Om vi summerar från vänster får vi 50 % + 34,1 % + 13,6 % = 97,7 %

Det betyder att vi ska täcka in alla värden från den lättaste vikten till en vikt som ligger 2 standardavvikelser över medelvärdet. Eftersom standardavvikelsen är 5 g så är 2 standardavvikelser 10 g. Det betyder det att medelvärdet måste vara 400 g + 10 g = 410 g.

Vi kontrollerar med räknaren att vi tänkt rätt! Vi använder verktyget RITA igen och ställer in så här:



Nu ritar vi och vi ser att vi får resultatet 97,7 %.



Denna uppgift var naturligtvis konstruerad så att man skulle använda sig av värdena för en och två standardavvikelser.

Vi ändrar kravet till att minst 90 % av burkarna ska innehålla minst 400 g tomater. Då kan vi inte direkt använda oss av kända värden.

### Lösa en ekvation

Nu vet vi inte medelvärdet och då blir det svårare. I alla verktyg vi använt har vi varit tvungna att mata in medelvärdet. Dags att använda räknarens **ekvationsverktyg**.

Tryck nu på tangenten  $\boxed{\text{math}}$  och gå med piltangenten ner till räknarens **Lösare**.

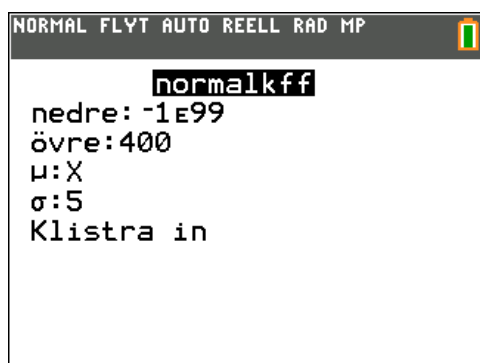


Tryck på  $\boxed{\text{enter}}$ . Nu ska vi mata in ekvationens vänster- och högerled E1 och E2.

I rutan för E1 skriver du in 0.10. Det motsvarar arean till vänster om den undre gränsen för normalfördelningen.

När du kommer till rutan E2 trycker du på  $\boxed{2\text{nd}}$   $\boxed{\text{distr}}$  då öppnas menyn för våra beräkningsverktyg. Välj nu 2: normalkff. Detta verktyg beräknar arean mellan två gränser i en normalfördelning med kända värden på medelvärde och standardavvikelse.

Nu öppnas en dialogruta där du fyller enligt nedan:



Den nedre gränsen kan ju vara hur litet negativt tal som helst eftersom normalfördelningar har oändlig utsträckning åt båda håll.  $-1E99$  eller  $10^{-99}$  är det minsta tal du kan skriva på räknaren.

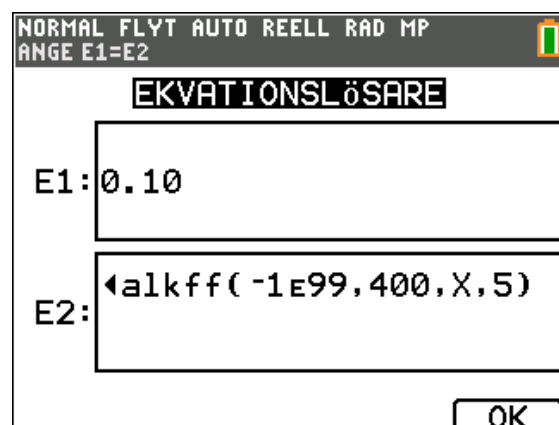
Den övre gränsen är ju 400 eftersom det är intervallet till vänster om som ska ha arean 0,10.

$\mu$  eller medelvärdet är vårt okända tal så vi skriver X där.

$\sigma$  eller standardavvikelsen är 5.

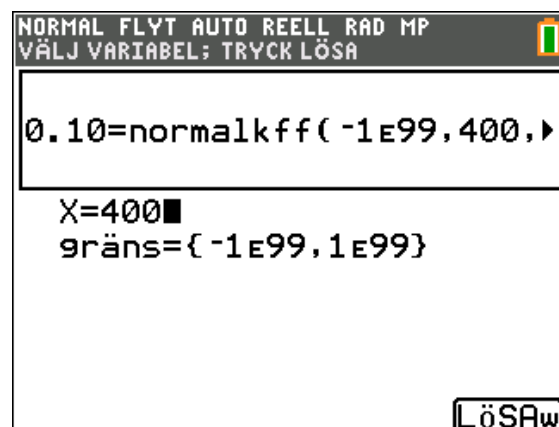
Markera nu Klistra in och tryck på  $\boxed{\text{enter}}$ .

Då ser det ut så här i ekvationslösaren:

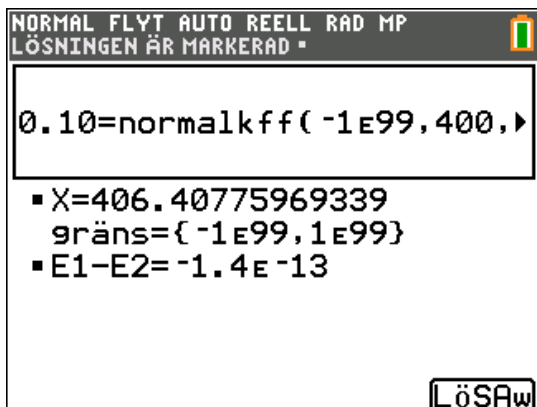


Vi ska alltså räkna ut för vilket värde på X som arean blir 0.10. Tryck nu på  $\boxed{\text{enter}}$ . Ekvationen står längst upp och sedan får du gissa ett värde på X. Vi gissar 400.

På sista raden står inom vilket intervall räknarens beräkningsverktyg ska leta efter lösningar.



Lår markören står kvar X=400 och tryck nu på  $\boxed{\alpha}$   $\boxed{\text{solve}}$ . Efter en kort stund kommer resultatet fram.

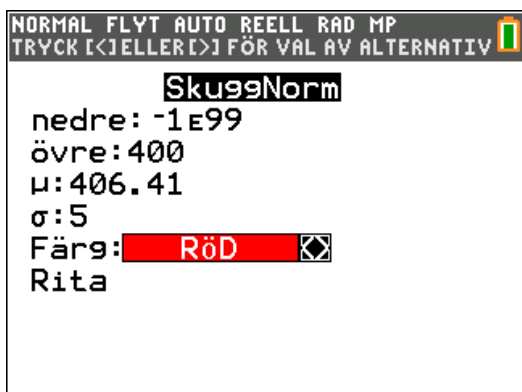


Medelvärde ska tydligen vara 406,4 för att 90 av burkarna ska ha en vikt som är minst 400 g.

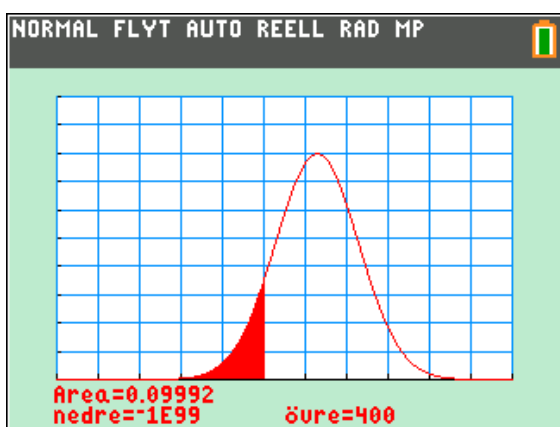
Vi kontrollerar att det stämmer med ett av verktygen som finns bland i menyn för fördelningar.

Tryck alltså  $2^{nd}$ [distr] välj RITA och sedan 1:SkuggaNorm.

Fyll nu i enligt nedan.  $\mu$  är ju medelvärdet som vi räknade ut.



Markera RITA och tryck på  $\text{enter}$ . Vi får att arean blir 0.09992 som avrundas till 0.10.



Med ekvationsverktyget kan du lösa problem både inom många områden. I en annan aktivitet tar vi upp hur man kan använda ekvationslösaren på olika sätt.

### Fördjupning

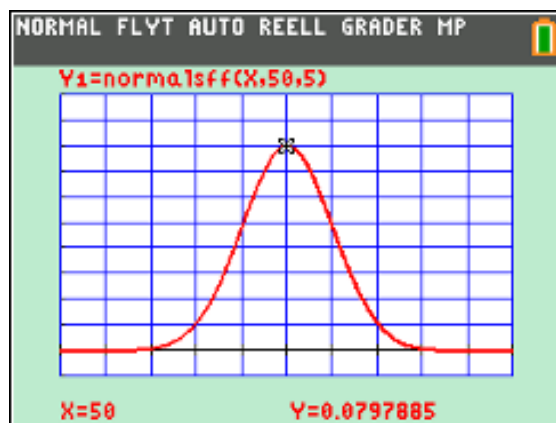
Hittills har vi använt räknarens inbyggda verktyg som man hittar genom att trycka på  $2^{nd}$ [distr]. Man kan också skriva in instruktioner för två av dem du använt, nämligen *normalsf* och *normalcdf* i editorn för inmatning av funktioner.

Vi tänker oss nu en normalfördelning där det uppskattade väntevärdet ( $\mu$ ) är 50 och standardavvikelsen ( $\sigma$ ) är 5. Det kan till exempel handla om vikten av en speciell sort av plommon.

Placera nu först markören vid Y1 i editorn för funktionsinmatning. Tryck sedan på  $2^{nd}$ [distr] och kopiera in normalsff. Fyll i mallen och skriv så att det blir så här i editorn för funktionsinmatning. (Tryck på  $\sqrt{\square}$ -tangenten).



Nu kan vi plotta denna funktion. med en bra fönsterinställning ser det ut så här i grafönstret.



Vi har alltså här en normalfördelningskurva med väntevärdet 50 och standardavvikelsen 5.



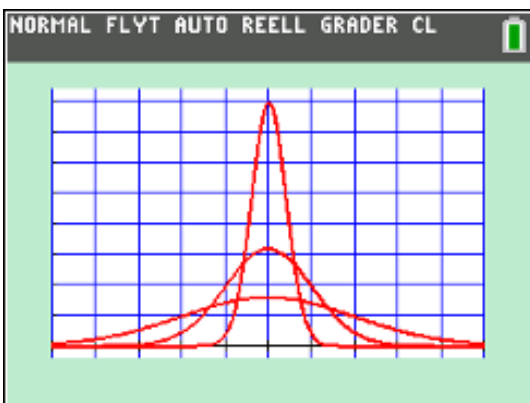
Denna funktion kan skrivas på ett allmänt sätt utan den inbyggda funktionen *normalssf*:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Konstanten *e* i formeln kommer du att arbeta mycket med i de senare kurserna. Det är en av matematikens, jämsides med  $\pi$ , viktigaste konstanter. Den utgör basen för s.k. naturliga logaritmer. Den har värdet 2,71828... Tryck på  $\boxed{2nd} \boxed{\div}$  så infogar du den i grundfönstret.

Du kan också få tre normalfördelningskurvor plottade samtidigt genom att infoga en lista med värdena 2,5, 5 och 10 i funktionsuttrycket. Arealen under varje kurva mot x-axeln är 1 för alla tre.

```
NORMAL FLYT AUTO REELL GRADER CL
Dia.91 Dia.92 Dia.93
Y1=normalssf(X,50,{2.5,10})
```



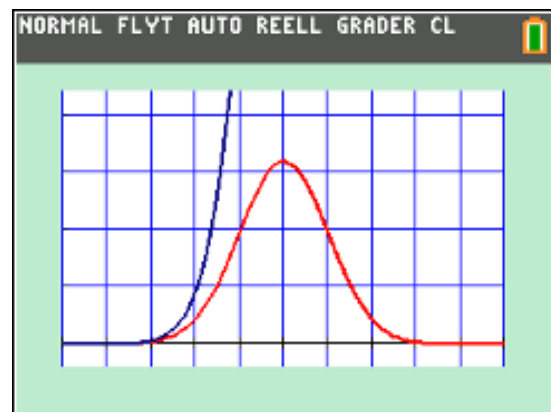
Nu kommer vi till det intressanta. Lägg nu in funktionen *normalkff* vid Y2. Placera då markören vid Y2 först och tryck sedan  $\boxed{2nd} \boxed{distr}$ . Välj där alternativ 2: *normalkff*. Fyll sedan i mallen så här.

```
NORMAL FLYT AUTO REELL GRADER CL
normalkff
nedre: -1E99
övre: X
μ: 50
σ: 5
Klistra in
```

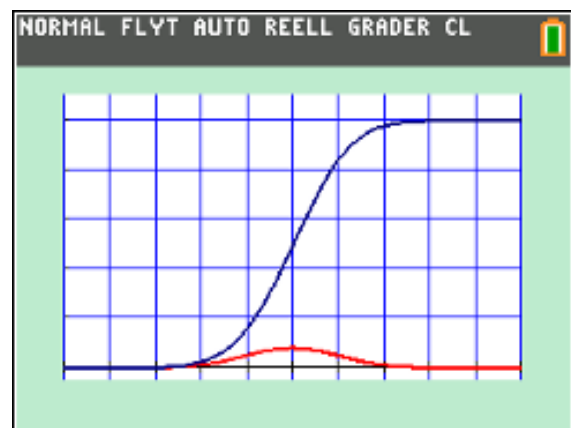
Infoga nu instruktionen i fönstret för funktionsinmatning genom att markera **Klistra in** och trycka på  $\boxed{enter}$ . Blir då så här:

```
NORMAL FLYT AUTO REELL GRADER CL
Dia.91 Dia.92 Dia.93
Y1=normalssf(X,50,5)
Y2=normalkff(-1E99,X,50,5)
Y3=
Y4=
Y5=
Y6=
Y7=
Y8=
```

Om vi nu plottar dessa två funktioner blir det så här:



Funktionen vid Y2 rusar iväg uppåt så här får vi ställa om fönstret. Den blå kurvan för Y2 visas då på ett bra sätt men den röda blir väldigt liten på höjden. Vi ställer nu om fönstret. Ymax är här 1.1.

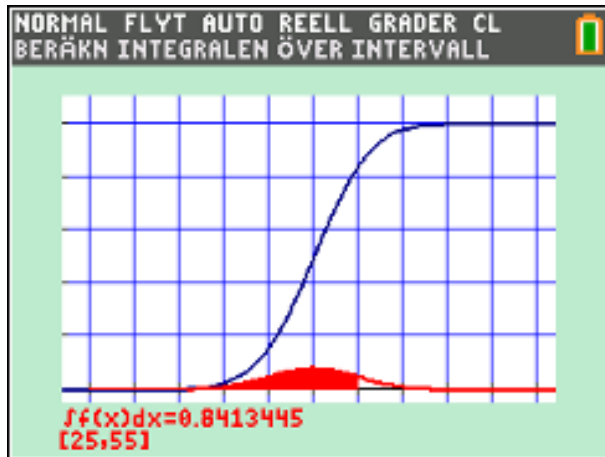


Låt oss nu säga att vi vill beräkna andelen plommon som har en vikt som är 55 g eller mindre.

Vi introducerar nu ett *nytt begrepp* som kommer upp i kurs 3 och 4. Här handlar det om att beräkna arean mellan två gränser för den röda kurvan. Det handlar här om begreppet *integraler*.

Tryck nu på  $\boxed{2nd}$   $\boxed{[calc]}$  och välj alternativ 7:  $\int f(x)dx$ .

Du får nu möjlighet att mata in två gränser. Välj nedre gräns 25 och övre gräns 55. Tryck sedan på  $\boxed{enter}$ . Nu fylls området under kurvan mellan gränserna ovan med rött och arean skrivs ut längst ner på skärmen. Vi får värdet 0,8413455.



Tryck på  $\boxed{trace}$  för en blå kurvan och skriv X=55 för spårningen. Ni visas värdet Y=0.8413447, dvs samma värde som markerade arean under den röda kurvan.

Den blå s-formade kurvan kallas *Fördelningsfunktionen* och anger sannolikheten att en normalfördelad variabel är mindre än eller lika med ett visst tal, i detta fall 55.

Normalfördelningen dyker också upp i andra sammanhang:

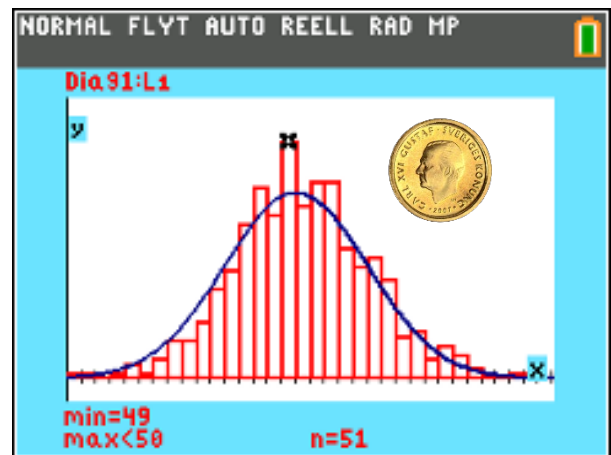
**Om en slant singlar till exempel 100 gånger kommer antal "krona" att vara ungefär binomialfördelat. Men eftersom varje slantsingling är oberoende av de övriga kommer summan att vara ungefär normalfördelad med väntevärdet 50.**

Nedan har vi gjort en simulering av försöket "kasta ett mynt 100 gånger och räkna antalet krona" och vi upprepar detta försök 500 gånger. Instruktionen blir *randbin*(antal, sannolikhet, antal försök) Talen i listan visar antalet krona vi får i varje försök.

L1	L2	L3	L4	L5	2
53	49.924	-----	-----	-----	
55	4.7669				
54	-----				
38					
48					
53					

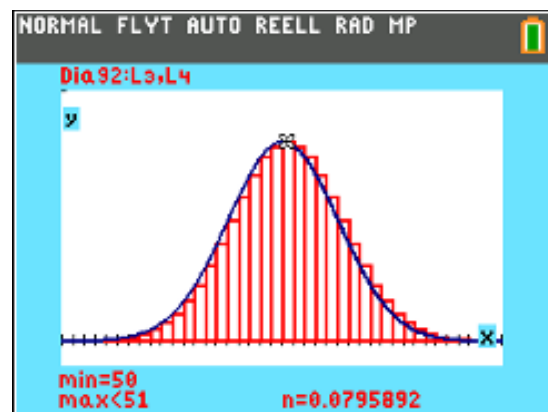
Vi har också beräknat medelvärde och standardavvikelse.

Om vi nu plottar våra värden i ett histogram där stapelbredden är 1 ser det ut enligt nedan. Vi har också lagt in en normalfördelningskurva med medelvärdet 50 och standardavvikelsen 5. Vi kan alltså approximera fördelningen i myntförsöket med en normalfördelning.

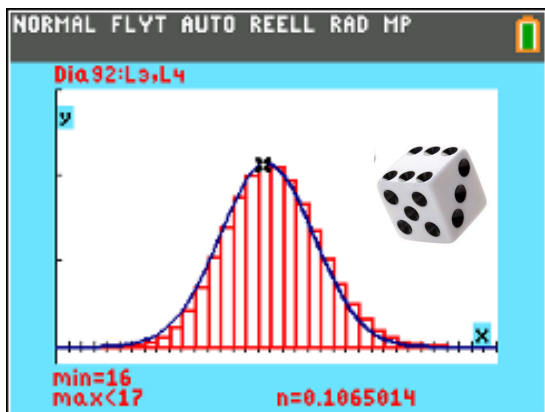


I plottningen nedan har vi räknat med de *exakta* sannolikheterna. Vi ser att staplarna nästan precis fyller ut arean under den blå normalfördelningskurvan. Sannolikheten att få precis 50 krona om man kastar ett mynt 100 gånger är ca 8 %.

Att använda normalfördelningen ger betydligt enklare beräkningar när man ska uppskatta olika sannolikheter.



Samma sak gäller om vi kastar en tärning och tittar på sannolikheterna att vi får en sexa ett visst antal gånger. Vi ser här att sannolikheten att få en sexa exakt 16 gånger på 100 kast är ca 10,7 %.



Ännu ett exempel: Om man kastar en tärning 5 gånger och tittar på *summan* av prickarna så närmar sig fördelningen också en normalfördelning. Man alstrar slumpalen (1, 2 ...6) i listorna L1 till och med L5. I L6 används formeln L1+L2+L3+L4+L5.

L2	L3	L4	L5	L6	σ
4	2	1	3	15	
6	3	6	6	23	
6	1	6	2	17	

