

# STATISTIKENS FRAMSTÄLLNING

## Normskördar

### Ämnesområde

Jord- och skogsbruk, fiske

### Statistikområde

Jordbrukets produktion

### Produktkod

JO0602

### Referenstid

Växtodlingsåret med skörd 2018

### Kontaktuppgifter

<b>Statistikansvarig myndighet</b>	Jordbruksverket
<b>Kontaktinformation</b>	Ann-Marie Karlsson
<b>E-post</b>	<a href="mailto:ann-marie.karlsson@jordbruksverket.se">ann-marie.karlsson@jordbruksverket.se</a>
<b>Telefon</b>	036-15 59 33

<b>Statistikproducent</b>	Statistiska centralbyrån, avdelningen för regioner och miljö, enheten för lantbruks- och energistatistik
<b>Kontaktinformation</b>	Gerda Ländell
<b>E-post</b>	<a href="mailto:gerda.landell@scb.se">gerda.landell@scb.se</a>
<b>Telefon</b>	010-479 68 07

## Innehåll

1	Statistikens sammanhang.....	3
2	Undersökningsdesign .....	3
2.1	Målstorheter .....	3
2.2	Ramförfarande .....	3
2.3	Förfaranden för urval och uteslutning .....	4
2.3.1	Urvalsförfarande.....	4
2.3.2	Uteslutning från insamling (cut-off) .....	4
2.4	Insamlingsförfarande.....	4
2.4.1	Datainsamlingsmetoder .....	4
2.4.2	Mätning.....	4
2.4.3	Bortfallsuppföljning .....	4
2.5	Bearbetningar.....	5
2.6	Granskning.....	5
2.6.1	Granskning under insamlingen .....	5
2.6.2	Granskning av mikrodata .....	6
2.6.3	Granskning av makrodata.....	6
2.6.4	Granskning av redovisning .....	6
2.7	Skattningsförfarande .....	6
2.7.1	Principer och antaganden .....	6
2.7.2	Skattningsförfarande för målstorheter.....	7
2.7.3	Skattningsförfarande för tillförlitlighet.....	8
2.7.4	Röjandekontroll .....	8
3	Genomförande .....	8
3.1	Kvantitativ information .....	8
3.2	Avvikelser från undersökningsdesignen .....	8

## 1 Statistikens sammanhang

Inom skördestatistiken framställs bland annat statistik om normala skördenivåer per hektar. Sedan början av 1960-talet har SCB årligen beräknat normskördar för de vanligaste spannmåls- och oljeväxtgrödorna samt för potatis och sockerbetor. Normskördarna baseras på underlag från alla odlingsformer som förekommer i landet och redovisas för riket, län, jordbrukets produktionsområden och skördeområden. Syftet med statistiken om normskördar är att visa den skörd som man kan förvänta sig under normala odlings- och väderbetingelser.

I detta dokument beskrivs framtagningen av statistiken om normskördar. Läs om statistikens kvalitet i kvalitetsdeklarationen (ersätter tidigare Beskrivning av statistiken), som åtföljer varje Statistiskt meddelande och finns tillgänglig på SCB:s och Jordbruksverkets webbplatser.

Statistiken är helt registerbaserad, det vill säga ingen insamling av data görs för detta ändamål, utan den bygger enbart på tidigare insamlad statistik. Information om underlagsundersökningarna ges i dokumenten ”Kvalitetsdeklaration” och ”Statistikens framställning” för Skörd av spannmål, trindsäd och oljeväxter, Skörd av potatis, Skörd av slåttervall och Skörd för ekologisk och konventionell odling. Mer information om äldre årgångar av dessa undersökningar finns i en så kallad Fördjupad dokumentation av statistiken. Den finns publicerad på Jordbruksverkets webbplats under ”Mer om statistiken, vegetabilieproduktion”.

## 2 Undersökningsdesign

### 2.1 Målstorheter

Inom denna skördestatistik redovisas hektarskördar i kilo per hektar av spannmål, trindsäd, oljeväxter, potatis, slåttervall och sockerbetor.

### 2.2 Ramförfarande

Normskördarna baseras på beräkningar som utgår från tidigare års skattningar av hektarskördar. Det finns således ingen egentlig undersökningspopulation av uppgiftslämnare för denna statistikprodukt. Uppgiftskällorna är undersökningarna ”Skörd av spannmål, trindsäd och oljeväxter” (produkt JO0601), ”Skörd av potatis” (produkt JO0603), ”Skörd av slåttervall” (produkt JO0606) respektive ”Skörd för ekologisk och konventionell odling” (produkt JO0608). För sockerbetor tas normskördarna fram utifrån leveransuppgifter baserade på information från samtliga sockerbetsodlare. Beräkningsunderlaget erhålls från Nordic Sugar.

Som urvalsram för skördeundersökningarna används Jordbruksverkets administrativa register för arealbaserade stöd, där jordbruksföretag som ansökt om arealersättning för innevarande år finns samlade.

Arealer vid jordbruksföretag som inte ansöker om arealbaserade stöd, och som därmed inte ingår i Jordbruksverkets administrativa register för arealbaserade stöd, ingår från år 2000 inte i skördeundersökningarna och därmed inte heller i beräkningsunderlaget för normskördarna. Detta bedöms dock för flertalet av de aktuella undersökningsgrödorna vara av marginell betydelse. I övrigt, se dokumenten för respektive årliga undersökning.

## **2.3 Förfaranden för urval och uteslutning**

### **2.3.1 Urvalsförfarande**

Ej tillämpligt. Normskördarna baseras på befintliga uppgifter om hektarskördar som SCB har tillgång till. Inget urval görs. Resultaten bygger på årliga skördedata från främst följande tre undersökningar:

- Skörd av spannmål, trindsäd och oljeväxter
- Skörd av potatis
- Skörd av slåttervall

Uppgifter om skörd av sockerbetor erhålls från uppköparen Nordic Sugar.

### **2.3.2 Uteslutning från insamling (cut-off)**

Den population som bildar underlag för skördestatistiken utgörs av jordbruksföretag med mer än 5,0 hektar åkermark. För sockerbetor finns dock ingen arealgräns utan alla jordbruksföretag med odling av sockerbetor för sockerproduktion ingår i underlaget för statistiken.

## **2.4 Insamlingsförfarande**

### **2.4.1 Datainsamlingsmetoder**

Se dokument för respektive årliga undersökning. Vid normskördeberäkningarna används befintliga uppgifter från skördeskattningarnas statistiska urvalsundersökningar. Beräkningsunderlaget finns vid SCB. De uppgifter som återanvänds är skattade hektarskördar och antalet observationer som varit underlag för skattningarna. Beräkningarna genomförs under försommaren när slutliga resultat gällande föregående års hektarskördar finns tillgängliga.

### **2.4.2 Mätning**

Se dokumentation för respektive årliga undersökning.

### **2.4.3 Bortfallsuppföljning**

Inget bortfall av beräkningsunderlag förekommer i denna undersökning.

## 2.5 Bearbetningar

I april publiceras slutliga skörderesultat avseende föregående år för spannmål, trindsäd, oljeväxter, potatis och slättervall för riket, län och produktionsområden. För normskördebehov beräknas därefter hektarskördar även för skördeområden. Statistiken läggs till de tids-serier som bildar underlag för normskördeberäkningarna.

För varje område och gröda beräknas en normskörd, om grödan odlas i tillräcklig omfattning så att resultat finns från skördeundersökningarna. Beräkningarna sker enligt två olika modeller beroende på antal år som grödan ingått i underlaget för skördestatistiken. Då tillräckligt antal år finns används en regressionsmodell. I övrigt används en medelvärdesmodell.

Normskörden i ett område, baserad på regressionsmodellen, beräknas som medeltalet av hektarskördarna enligt skördeskattningarna (leveransuppgifter för sockerbetor) under de senaste 15 åren före det aktuella normskördeåret plus en beräknad skördeutveckling från 15-årsperiodens mitt till och med det aktuella skördeåret. Antalet jordbruksföretag med skördeuppgift måste vara minst fem för varje enskilt år som ingår i beräkningsunderlaget. Antalet år med tillräckligt antal jordbruksföretag måste vara minst tio. Vid beräkningen används en regressionsmodell med årtal som förklarande variabel och en stabiliseringsvariabel. Skördeutvecklingen beskrivs med en rät linje, den så kallade trendlinjen.

I medelvärdesmodellen beräknas avkastningen som medelvärden av de senaste 10 årens hektarskördar frånsett den lägsta och den högsta hektarskörderna under perioden. Metoden benämns "trimmat tioårsmedelvärde". Antalet jordbruksföretag med skördeuppgift måste vara minst fem för varje enskilt år som ingår i underlaget. Antalet år med tillräckligt antal jordbruksföretag måste vara minst tio.

Vid beräkningen av femårsmedelvärden måste antalet jordbruksföretag med skördeuppgift vara minst tio för varje enskilt år av de fem år som ingår i beräkningsunderlaget.

Imputering görs inte varken vid partiella bortfall eller vid objektsbortfall.

## 2.6 Granskning

Resultaten av normskördeberäkningarna granskas av lantbrukskunniga personer vid SCB.

### 2.6.1 Granskning under insamlingen

Inom de undersökningar som utgör underlag för beräkningarna finns maskinella kvalitetskontroller i webbsystemet som vägleder uppgiftslämnarna. Under hela insamlingsperioden sker både maskinell och

manuell granskning. Se dokumentationen för respektive ingående undersökning för mer detaljerad information.

### 2.6.2 Granskning av mikrodata

Se dokumentation för respektive årliga undersökning – där det framgår att granskning av mikrodata sker av lantbrukskunnig personal i lantbrukssystemet genom scatter-plot-diagram.

### 2.6.3 Granskning av makrodata

Skattningarna av normskördar på riksnivå jämförs med motsvarande statistik från tidigare år. Granskning normskörderesultaten för skördeområden görs med hjälp av att resultaten redovisas på kartor, varvid rimlighetsbedömningar görs utifrån det geografiska läget.

### 2.6.4 Granskning av redovisning

Inför publicering kontrolleras de saklogiska samband som ska råda mellan värden i tabellerna. Korrekturläsning genomförs av någon som känner ämnesområdet men är fristående från produktionsarbetet.

Genom ett gott samarbete mellan Jordbruksverket och SCB kan skördestatistiken publiceras på båda myndigheternas webbplatser – till nytta för användarna.

## 2.7 Skattningsförfarande

### 2.7.1 Principer och antaganden

Målstorhet i undersökningen är den så kallade normskörden och den beräknas per gröda uppdelat på områden (riket, län, produktionsområden och skördeområden).

Utgångspunkt för beräkningarna är en multipel linjär regressionsmodell baserad på de senaste årens skördeuppskattningar. Regressionsmodellen kräver att 21 års data finns tillgänglig. För vissa grödor där tillräckligt många års data saknas, beräknas normskördar istället utifrån medelvärden baserade på 5- respektive 10-års data.

Regressionsmodellen ges av sambandet:

$$H_t = \alpha + \beta \times t + \gamma \times S_t + \varepsilon_t \quad (1)$$

Där  $t$  betecknar tiden mätt i år,  $H_t$  betecknar den skattade hektarskörden år  $t$ , och där variabeln  $S_t$  definieras som

$$S_t = H_t - \frac{1}{6} \sum_{k=1}^6 H_{t-k}, \quad (2)$$

och alltså utgör skillnaden mellan hektarskörden för år  $t$  och medelhektarskörden för de sex närmast föregående åren. Stabiliseringsvariabeln  $S_t$  har införts i syfte att stabilisera normskördeutvecklingen. Residualtermen  $\varepsilon_t$  i sambandet (1) antas vara normalfördelad med

väntevärde noll och konstant varians. Residualtermer för olika år antas vidare vara oberoende av varandra. Modellens tre parametrar ( $\alpha$ ,  $\beta$ , och  $\gamma$ ) skattas med minsta-kvadrat-metoden.

Regressionsperioden ges av de senaste 15 åren före det aktuella skördeåret. Införandet av stabiliseringsvariabeln kräver dock att data från ytterligare 6 år finns tillgängliga.

För de grödor där normskördar beräknas utifrån medelvärden görs inga modellantaganden.

### 2.7.2 Skattningsförfarande för målstorheter

Normskörden för det aktuella skördeåret ( $t+1$ ) betecknas här  $N_{t+1}$ . När regressionsmodellen (1) används beräknas normskörden enligt uttrycket

$$N_{t+1} = \hat{\alpha} + \hat{\beta} \times (t + 1) + \hat{\gamma} \times \bar{S}_t, \quad (3)$$

där  $\hat{\alpha}$ ,  $\hat{\beta}$ , och  $\hat{\gamma}$  betecknar minsta-kvadrat skattningarna av modellens parametrar, och  $\bar{S}_t$  betecknar medelvärdet av stabiliseringsvariabeln över den 15 år långa regressionsperioden. Normskörden kan i detta fall även uttryckas som

$$N_{t+1} = \bar{H} + \hat{\beta}(t + 1 - \bar{t}), \quad (4)$$

där  $\bar{H}$  och  $\bar{t}$  betecknar medelvärdena över regressionsperioden för de skattade hektarskördarna  $H_t$  samt tidsvariabeln  $t$ .

För grödor med mindre än 21 år av tillgänglig data beräknas normskörden som medelvärden över skattade hektarskördar. När det finns 10 år av tillgänglig data beräknas normskörden utifrån så kallade trimmade medelvärden där de minsta och största hektarskördarna undantas från medelvärdet. Förfarandet används för att minska variationen mellan år och ge en mer stabil normskörd. Normskördar baserade på trimmade 10-års medelvärden kan uttryckas som

$$N_{t+1} = \frac{1}{8} \sum_{k=0}^9 (H_{t-k} - H_{max} - H_{min}), \quad (5)$$

där

$$H_{max} = \max(H_t, \dots, H_{t-9})$$

och

$$H_{min} = \min(H_t, \dots, H_{t-9}).$$

Slutligen, för grödor med färre än 10 år, och minst 5 år av tillgänglig data, beräknas normskörden som ett vanligt 5-års medelvärde enligt:

$$N_{t+1} = \frac{1}{5} \sum_{k=0}^4 H_{t-k} \quad (6)$$

### **2.7.3 Skattningsförfarande för tillförlitlighet**

Osäkerhetsmått i form av medelfel redovisas inte. För att i någon mån indikera resultatens säkerhet redovisas från och med 2007 antalet jordbruksföretag vars hektarskörda utgjort beräkningsunderlaget för de år som ingått vid framtagningen av grödans normskörd respektive medelvärde i det aktuella området. Läs mer om statistikens tillförlitlighet i kvalitetsdeklarationen (för tidigare årgångar Beskrivning av statistiken).

### **2.7.4 Röjandekontroll**

För att normskörd, trimmat tioårsmedelvärde eller femårsmedelvärde för en gröda ska redovisas krävs att det finns uppgifter från allra minst 5 företag för minst 21, 10 respektive 5 år. Därmed finns ingen röjanderisk.

## **3 Genomförande**

Det har skett en dialog samt en årlig uppföljning av statistikens inriktning med de viktigaste användarna genom dels de användarråd som Jordbruksverket anordnar, dels via kontakter med andra statistik användare. Inga större omläggningar har föreslagits. Den trendlinje som tas fram inom regressionsmodellen baseras på jämförelser av skördeutfallet under en följd av år. Därmed är det angeläget att uppläggningsen av skördeundersökningarna inte förändras på något sätt som orsakar tidsseriebrott i skördeutvecklingen, utan att de genomförs på samma sätt år efter år.

### **3.1 Kvantitativ information**

Se dokument för respektive årliga undersökning.

### **3.2 Avvikelser från undersökningsdesignen**

Undersökningen har genomförts som planerat. Inga avvikelser har gjorts.