



Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

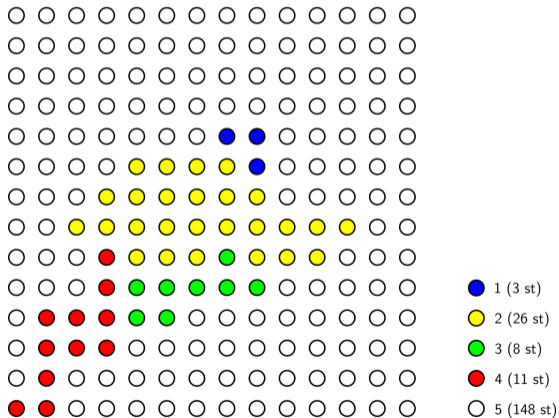
Institutionen för skoglig resurshushållning

Skattning och variansskattning för ny design

Anton Grafström, anton.grafstrom@slu.se

Exempeltrakt

Klassning av ytor på trakt



Traktens värde

Vid skattning av areal av klass k är traktens värde andelen av arealen av trakten som har klass k . Alltså, om $Y(x)$ är värdet för en trakt på koordinaten x , då är traktens värde

$$Y(x) = a_k/a_T,$$

där a_T är traktens totala areal (dvs. 196 gånger provytans areal a) och a_k är totala arealen av klass k på trakten. Om $a_k(i)$ är areal av klass k på provyta i , så

$$Y(x) = \frac{1}{a_T} \sum_{i=1}^{196} a_k(i).$$

Urval för skattning av traktens värde

På trakten finns $N = 196$ provytor som utgör populationen som samplas. Provytorna klassas till H strata, vilket ger antalet inom varje klass N_1, N_2, \dots, N_H . Enligt algoritmen bestäms stickprovsstorlekarna n_1, n_2, \dots, n_H för respektive stratum. Om $h(i)$ är stratum för provyta i , ges inklusionssannolikheten för provyta i av

$$\pi_i = \frac{n_{h(i)}}{N_{h(i)}}.$$

I exemplet har vi $H = 5$ klasser, samt $N_1 = 3, N_2 = 26, N_3 = 8, N_4 = 11$ och $N_5 = 148$. Säg att $n_1 = n_2 = n_3 = 3$ och $n_4 = 1$ samt $n_5 = 0$.

Inklusionssannolikheter på trakt

För exemplet ger detta inklusionssannolikheterna för provytorna på trakten

Strata	Inklusionssannolikhet
1	$n_1/N_1 = 3/3 = 1$
2	$n_2/N_2 = 3/26$
3	$n_3/N_3 = 3/8$
4	$n_4/N_4 = 1/11$
5	$n_5/N_5 = 0/148 = 0$



Vi skattar traktens värde

För skattning av traktens värde används vanlig HT-skattning

$$\hat{Y}(x) = \frac{1}{a_T} \sum_{i \in S(x)} \frac{a_k(i)}{\pi_i},$$

där summan är över de valda provytorna på trakten, och $a_k(i)$ är arealen av provyta i som täcks av klass k .

Vi skattar traktens värde (exempel)

Säg att vi i exemplet observerade 25% av 1 yta i klass k i stratum 1, dvs areal $0.25a$ där a är provytans areal. Vi observerar också 90% av 1 yta i klass k i stratum 2 ($0.9a$). Då skattar vi traktens värde för klass k som

$$\begin{aligned}\hat{Y}(x) &= \frac{1}{a_T} \left(\frac{0.25a}{3/3} + \frac{0.9a}{3/26} \right) = \frac{1}{196a} \left(\frac{0.25a}{3/3} + \frac{0.9a}{3/26} \right) = \frac{1}{196} \left(\frac{0.25}{3/3} + \frac{0.9}{3/26} \right) \\ &= (0.25 + 7.8)/196 = 0.04107143,\end{aligned}$$

Vi skattar alltså att drygt 4% av trakten täcks av klass k .

Skattning av total för ett område

Låt urvalsramen ha areal a_F (inklusive buffer). Om n trakter har valts inom ramen (för den valda tätheten på stickprovet), då är samplingintensiteten

$$\pi(x) = \frac{n}{a_F}.$$

Skattning av total på ett område blir då

$$\hat{Y} = \sum_{x \in S} \frac{\hat{Y}(x)}{\pi(x)} = \frac{a_F}{n} \sum_{x \in S} \hat{Y}(x),$$

där S är alla valda traktcentrum/trakter (över hela Sverige). $\hat{Y}(x)$ är det skattade värdet för trakten på koordinat x . Vid beräkning av traktens värde räknas bara det som är inom aktuellt område (men delas med hela traktens areal).

Alternativ för variansskattning, oberoende trakter

$$\hat{V}_{iid}(\hat{Y}) = a_F^2 \frac{S_Y^2}{n},$$

där

$$S_Y^2 = \frac{1}{n-1} \sum_{x \in S} \left(\hat{Y}(x) - \frac{1}{n} \sum_{x \in S} \hat{Y}(x) \right)^2.$$

Alternativ för variansskattning, rumslig balans

$$\hat{V}_{SB}(\hat{Y}) = \frac{a_F^2}{n^2} \sum_{x \in S} \frac{n_x}{n_x - 1} \left(\hat{Y}(x) - \frac{1}{n_x} \sum_{x' \in D_x} \hat{Y}(x') \right)^2,$$

där D_x innehåller koordinaterna för de n_x mest lika trakterna till dev vid x (även trakten vid x). Likhet mäts i hjälpvariabler som vid urval. Storleken n_x bör vara ganska liten och normalt inte överstiga 4. Om storleken är $n_x = 2$ får vi

$$\hat{V}_{SB}(\hat{Y}) = \frac{a_F^2}{2n^2} \sum_{x \in S} \left(\hat{Y}(x) - \hat{Y}(x') \right)^2,$$

där x' är koordinaten för den trakt som är mest lik trakten på koordinaten x , mätt i hjälpvariablerna som använts vid urvalet.

Om vi istället vill skatta totalt antal plantor eller total volym, då ska traktens värde vara traktens totala antal/volym delat på traktens areal. Detta värde skattas på samma sätt som areal av en klass. T.ex. låt u_i vara antal/volym på provyta i . Då är traktens värde

$$Y(x) = \frac{1}{a_T} \sum_{i=1}^{196} u_i,$$

och det skattas med

$$\hat{Y}(x) = \frac{1}{a_T} \sum_{i \in S(x)} \frac{u_i}{\pi_i},$$

Enkel analys

- Trots (eller kanske snarare tack vare) komplicerad design får vi väldigt enkel analys. Det bör underlätta användandet av data från den nya designen. Alla skattningar kan göras på samma sätt över hela landet och för regioner eller län, bara målvariabeln ändras.
- Varje provyta bör förses med indikatorer för region och län mm. Då behövs endast en algoritm för beräkningar av skattningar och variansskattningar.
- Detta bör implementeras i ett R-paket som också innehåller all designinformation. Då kan data hämtas in (från webportal) och skattningar blir enkelt med en manual.
- Även ett webbaserat gränssnitt för skattningar blir enkelt.