

GRTS Survey Designs for an Area Resource

Thomas Kincaid

May 23, 2011

Contents

1 Preliminaries	1
2 Shapefile attribute data	2
3 Unstratified, equal probability, GRTS survey design	4
4 Unstratified, unequal probability, GRTS survey design	5
5 Stratified, equal probability, GRTS survey design	7
6 Unstratified, unequal probability, GRTS survey design with an oversample and a panel structure for survey over time	9

This document presents example GRTS survey designs for an area resource. The area resource used in the designs is Omernik level 3 ecoregions within Utah. Four survey designs will be presented: (1) an unstratified, equal probability design; (2) an unstratified, unequal probability design; (3) a stratified, equal probability design; and (4) an unstratified, unequal probability design with an oversample and a panel structure for survey over time. The sampling frame used for the survey designs is contained in either an ESRI shapefile or an `sp` package object. The frame contains the coordinates for a set of polygons that define the area resource in addition to attribute data associated with the polygons. The coordinate system for the set of points in the sampling frame is an equal area projection rather than latitude and longitude. An equal area projection is used so that calculation of distance between points is valid.

1 Preliminaries

The initial step is to use the library function to load the `spsurvey` package. After the package is loaded, a message is printed to the R console indicating that the `spsurvey` package was loaded successfully.

Load the `spsurvey` package

```
> library(spsurvey)
```

Version 2.1 of the `spsurvey` package was loaded successfully.

2 Shapefile attribute data

The next step is to read the attribute data from the shapefile. The `read.dbf` function in the `spsurvey` package is used to read the attribute (dbf) file in the shapefile and assign it to a data frame named `att`. The initial six lines in the `att` data frame are printed using the `head` function.

The ecoregion attribute will be used to define stratum codes and unequal selection probability (multidensity) categories for the survey designs. Ecoregion is contained in a variable named "level3_nam" and includes seven unique values. Frame area is summarized for the ecoregion attribute. Note that ecoregion area measured in hectares is contained in the variable named "area_ha". The `tapply` function is used to calculate total area for each ecoregion. The `addmargins` function is applied to the output from `tapply` to calculate total area for all all ecoregions, and the `round` function is used to round value to whole numbers. Finally, the resulting table is displayed.

Read the attribute table from the shapefile

```
> att <- read.dbf("eco_l3_ut")
```

Display the initial six lines in the attribute data frame

```
> head(att)
```

	level3	level3_nam	area_ha	area_mdm
1	80	Northern Basin and Range	263999.04	2639990439
2	18	Wyoming Basin	291058.83	2910588302
3	13	Central Basin and Range	8206454.64	82064546355
4	19	Wasatch and Uinta Mountains	4256949.15	42569491524
5	20	Colorado Plateaus	8579716.31	85797163092
6	21	Southern Rockies	54090.91	540909129

Summarize frame area by ecoregion

```
> temp <- tapply(att$area_ha, att$level3_nam, sum)
> temp <- round(addmargins(temp), 0)
> temp
```

Central Basin and Range	Colorado Plateaus
8206455	8579716
Mojave Basin and Range	Northern Basin and Range
193186	263999
Southern Rockies	Wasatch and Uinta Mountains
94644	4353925
Wyoming Basin	Sum
291059	21982984

Ecoregions in Utah are displayed in Figure 1. To produce the figure, first the `read.shape` function in the `spsurvey` package is used to read the shapefile and assign it to an object named `shp`. The `shp` object takes the form of a spatial data object defined in the `sp` package. Specifically, `shp` belongs to class "SpatialPolygonsDataFrame". For further information about spatial data objects, see documentation for the `sp` package. The `spplot` function in the `sp` package is used to create the figure.

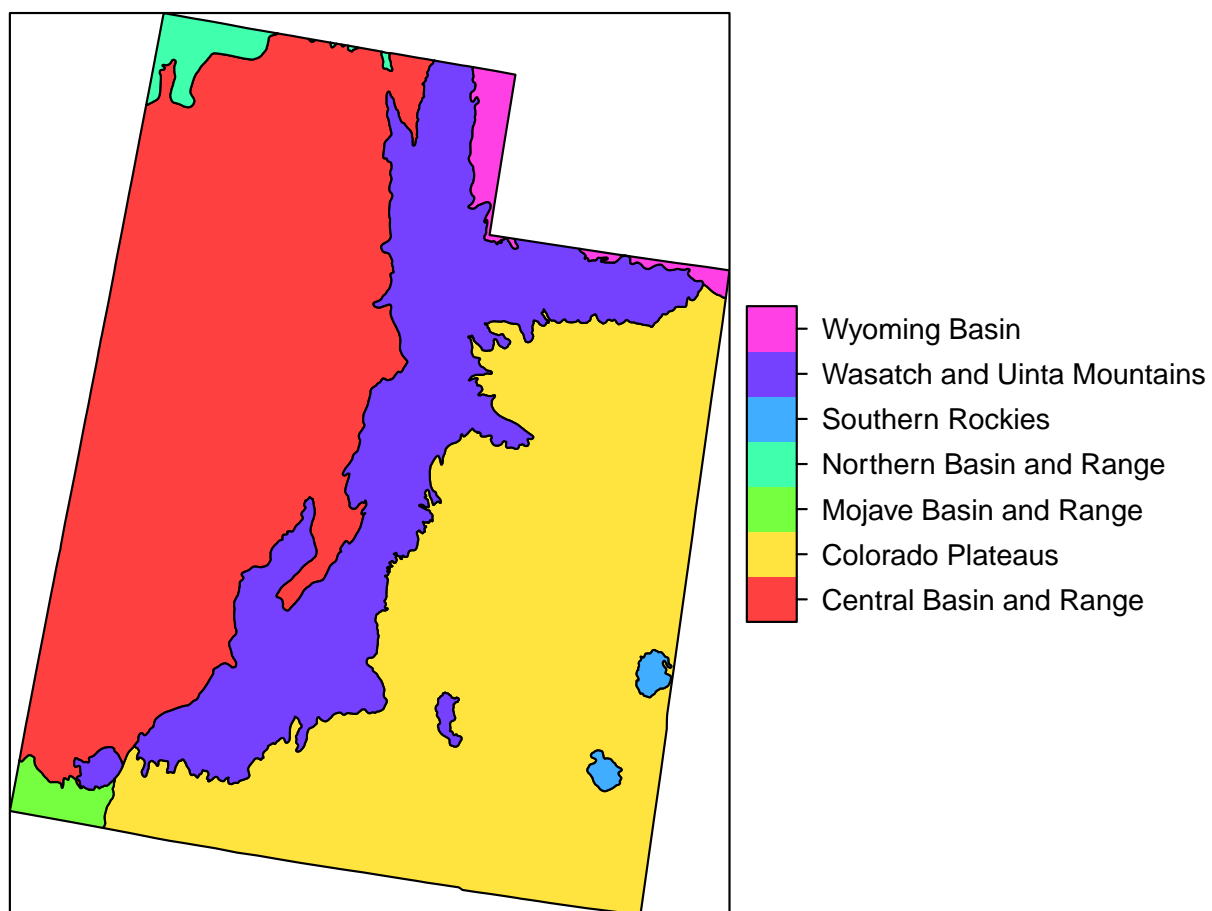


Figure 1: Ecoregions in Utah.

3 Unstratified, equal probability, GRTS survey design

The first survey design is an unstratified, equal probability design. The `set.seed` function is called so that, if necessary, the designs can be replicated.

The initial step is to create a list named `Equaldsgn` that contains information for specifying the survey design. Since the survey design is unstratified, the list contains a single item named "None" that also is a list. The "None" list includes two items: `panel`, which is used to specify the sample size for each panel, and `seltype`, which is used to input the type of random selection for the design. For this example, `panel` is assigned a single value named "PanelOne" that is set equal to 115, and `seltype` is assigned the value "Equal", which indicates equal probability selection.

The `grts` function in the `spsurvey` package is called to select the survey design. The following arguments are included in the call to `grts`: (1) `design`: the named list of stratum design specifications, which is assigned the `Equaldsgn` list; (2) `DesignID`: name for the design, which is used to create a site ID for each site and is assigned the value "EQUAL"; (3) `type.frame`: the type of frame, which is assigned the value "area" to indicate an area resource; (4) `src.frame`: source of the frame, which is assigned the value "shapefile" to indicate a shapefile frame; (5) `in.shape`: name of the input shapefile, which is assigned the value "eco_l3_ut"; (6) `att.frame`: the data frame of attributes associated with elements in the frame, which is assigned the `att` data frame; and (7) `shapefile`: option to create a shapefile containing the survey design information, which is assigned `FALSE`.

During execution of the `grts` function, messages are printed that indicate the initial number of hierarchical levels used for the GRTS grid, the current number of levels, and the final number of levels. The set of messages is printed for each stratum, and is labeled with the stratum name. For this example, the set of messages is labeled "None", i.e., the name used in the `Equaldsgn` list. Upon completion of the call to `grts`, the initial six sites for the survey design and a design summary are printed.

Call the `set.seed` function so that the design can be replicated

```
> set.seed(4447864)
```

Create the design list

```
> Equaldsgn <- list(None = list(panel = c(PanelOne = 115), seltype = "Equal"))
```

Select the sample

```
> Equalsites <- grts(design=Equaldsgn,
+                   DesignID="EQUAL",
+                   type.frame="area",
+                   src.frame="shapefile",
+                   in.shape="eco_l3_ut",
+                   att.frame=att,
+                   shapefile=FALSE)
```

Stratum: None

Initial number of levels: 4

Current number of levels: 4

Final number of levels: 4

Print the initial six lines of the survey design

```
> head(Equalsites@data)
```

	siteID	xcoord	ycoord	mdcaty	wgt	stratum	panel	EvalStatus
1	EQUAL-001	-1361882	1986078	Equal	1911563856	None	PanelOne	NotEval
2	EQUAL-002	-1304045	2195914	Equal	1911563856	None	PanelOne	NotEval
3	EQUAL-003	-1269657	1803336	Equal	1911563856	None	PanelOne	NotEval
4	EQUAL-004	-1311248	1689394	Equal	1911563856	None	PanelOne	NotEval
5	EQUAL-005	-1353690	2074387	Equal	1911563856	None	PanelOne	NotEval
6	EQUAL-006	-1389763	2190386	Equal	1911563856	None	PanelOne	NotEval

	EvalReason	level3	level3_nam	area_ha
1		13	Central Basin and Range	8206455
2		13	Central Basin and Range	8206455
3		20	Colorado Plateaus	8579716
4		20	Colorado Plateaus	8579716
5		13	Central Basin and Range	8206455
6		13	Central Basin and Range	8206455

Print the survey design summary

```
> dsgnsum(Equalsites)
```

Design Summary: Number of Sites

```
stratum
None Sum
115 115
```

4 Unstratified, unequal probability, GRTS survey design

The second survey design is an unstratified, unequal probability design. Ecoregions are used to identify multidensity categories. List Unequaldsgn is assigned design specifications. Since the survey design is unstratified, Unequaldsgn includes a single list named "None" that contains three items: panel, seltype, and caty.n. The value for panel is the same as for the equal probability design, and seltype is assigned "Unequal" to indicate unequal selection probabilities. The third item, caty.n, assigns sample sizes for each of seven multidensity categories, where ecoregion names are used as the categories. Note that the sum of sample sizes provided in caty.n must equal the value in panel.

For this survey design, a shapefile will be used as the sampling frame. The following arguments are included in the call to grts: (1) design: assigned the Unequaldsgn list; (2) DesignID: assigned the value "UNEQUAL"; (3) type.frame: assigned the value "area"; (4) src.frame: assigned the value "shapefile"; (5) in.shape: assigned the value "eco_l3_ut"; (6) att.frame: assigned the att data frame; (7) mdcaty: name of the column in the attributes data frame that identifies the unequal probability category for each element in the frame, which is assigned the value "level3_nam"; and (8) shapefile: assigned the value FALSE. Upon completion of the call to grts, the initial six sites for the survey design and a design summary are printed.

Create the design list

```
> Unequaldsgn <- list(None=list(panel=c(PanelOne=115),
+                               seltype="Unequal",
```

```

+                               caty.n=c("Central Basin and Range"=25,
+                                         "Colorado Plateaus"=25,
+                                         "Mojave Basin and Range"=10,
+                                         "Northern Basin and Range"=10,
+                                         "Southern Rockies"=10,
+                                         "Wasatch and Uinta Mountains"=25,
+                                         "Wyoming Basin"=10)))

```

Select the sample

```

> Unequalsites <- grts(design=Unequaldsgn,
+                       DesignID="UNEQUAL",
+                       type.frame="area",
+                       src.frame="shapefile",
+                       in.shape="eco_l3_ut",
+                       att.frame=att,
+                       mdcaty="level3_nam",
+                       shapefile=FALSE)

```

Stratum: None

Initial number of levels: 4

Current number of levels: 4

Current number of levels: 6

Current number of levels: 7

Final number of levels: 7

Print the initial six lines of the survey design

```
> head(Unequalsites@data)
```

	siteID	xcoord	ycoord	mdcaty	wgt	stratum
1	UNEQUAL-001	-1250114	1975757	Wasatch and Uinta Mountains	1741570045	None
2	UNEQUAL-002	-1233142	1804211	Colorado Plateaus	3431886524	None
3	UNEQUAL-003	-1308234	2197419	Central Basin and Range	3282581854	None
4	UNEQUAL-004	-1408362	1899295	Central Basin and Range	3282581854	None
5	UNEQUAL-005	-1200400	1849394	Colorado Plateaus	3431886524	None
6	UNEQUAL-006	-1177572	1873275	Colorado Plateaus	3431886524	None

	panel	EvalStatus	EvalReason	level3	area_ha
1	PanelOne	NotEval		19	4256949
2	PanelOne	NotEval		20	8579716
3	PanelOne	NotEval		13	8206455
4	PanelOne	NotEval		13	8206455
5	PanelOne	NotEval		20	8579716
6	PanelOne	NotEval		20	8579716

Print the survey design summary

```
> dsgnsum(Unequalsites)
```

Design Summary: Number of Sites Classified by mdcaty (Multidensity Category)

mdcaty	
Central Basin and Range	Colorado Plateaus
30	23
Mojave Basin and Range	Northern Basin and Range
11	9
Southern Rockies	Wasatch and Uinta Mountains
11	21
Wyoming Basin	Sum
10	115

5 Stratified, equal probability, GRTS survey design

The third survey design is a stratified, equal probability design. Ecoregions are used to identify strata. List Stratdsgn is assigned design specifications. The ecoregion attribute is used to identify strata. Stratdsgn includes seven lists, one for each stratum. The names for the lists match the levels of the stratum variable, i.e., the unique values of the ecoregion attribute. Each list in Stratdsgn contains two items: panel and seltype. The value for panel is the same as for the equal probability design, and seltype is assigned "Equal".

For this survey design, an sp package object will be used as the sampling frame. Recall that the read.shape function was used to read the shapefile and assign its output to an sp object named shp. The following arguments are included in the call to grts: (1) design: assigned the Stratdsgn list; (2) DesignID: assigned the value "STRATIFIED"; (3) type.frame: assigned the value "area"; (4) src.frame: assigned the value "sp.object" to indicate that the sampling frame is provided by an sp object; (5) sp.object: name of the sp object, which is assigned the shp object; (6) att.frame: assigned the att data frame; (7) stratum: name of the column in the attributes data frame that identifies the stratum code for each element in the frame, which is assigned the value "level3_nam"; and (8) shapefile: assigned the value FALSE. Upon completion of the call to grts, the initial six sites for the survey design and a design summary are printed.

Create the design list

```
> Stratdsgn <- list("Central Basin and Range"=list(panel=c(PanelOne=25),
+
+                                     seltype="Equal"),
+
+       "Colorado Plateaus"=list(panel=c(PanelOne=25),
+                                     seltype="Equal"),
+
+       "Mojave Basin and Range"=list(panel=c(PanelOne=10),
+                                     seltype="Equal"),
+
+       "Northern Basin and Range"=list(panel=c(PanelOne=10),
+                                     seltype="Equal"),
+
+       "Southern Rockies"=list(panel=c(PanelOne=10),
+                                     seltype="Equal"),
+
+       "Wasatch and Uinta Mountains"=list(panel=c(PanelOne=25),
+                                     seltype="Equal"),
+
+       "Wyoming Basin"=list(panel=c(PanelOne=10),
+                                     seltype="Equal"))
```

Select the sample

```
> Stratsites <- grts(design=Stratdsgr,
+                     DesignID="STRATIFIED",
+                     type.frame="area",
+                     src.frame="sp.object",
+                     sp.object=shp,
+                     att.frame=att,
+                     stratum="level3_nam",
+                     shapefile=FALSE)
```

Stratum: Central Basin and Range

Initial number of levels: 3

Current number of levels: 3

Current number of levels: 4

Final number of levels: 4

Stratum: Colorado Plateaus

Initial number of levels: 3

Current number of levels: 3

Current number of levels: 4

Final number of levels: 4

Stratum: Mojave Basin and Range

Initial number of levels: 2

Current number of levels: 2

Current number of levels: 3

Final number of levels: 3

Stratum: Northern Basin and Range

Initial number of levels: 2

Current number of levels: 2

Current number of levels: 4

Final number of levels: 4

Stratum: Southern Rockies

Initial number of levels: 2

Current number of levels: 2

Current number of levels: 4

Final number of levels: 4

Stratum: Wasatch and Uinta Mountains

Initial number of levels: 3

Current number of levels: 3

Current number of levels: 4

Final number of levels: 4

Stratum: Wyoming Basin

Initial number of levels: 2

Current number of levels: 2


```
Current number of levels: 3
Current number of levels: 4
Final number of levels: 4
```

Print the initial six lines of the survey design

```
> head(Stratsites@data)
```

	siteID	xcoord	ycoord	mdcaty	wgt	stratum
1	STRATIFIED-001	-1332619	2220346	Equal	3282581854	Central Basin and Range
2	STRATIFIED-002	-1491714	2108847	Equal	3282581854	Central Basin and Range
3	STRATIFIED-003	-1359436	2108766	Equal	3282581854	Central Basin and Range
4	STRATIFIED-004	-1559161	1745238	Equal	3282581854	Central Basin and Range
5	STRATIFIED-005	-1500699	2000075	Equal	3282581854	Central Basin and Range
6	STRATIFIED-006	-1395238	2010544	Equal	3282581854	Central Basin and Range

	panel	EvalStatus	EvalReason	level3	area_ha
1	PanelOne	NotEval		13	8206455
2	PanelOne	NotEval		13	8206455
3	PanelOne	NotEval		13	8206455
4	PanelOne	NotEval		13	8206455
5	PanelOne	NotEval		13	8206455
6	PanelOne	NotEval		13	8206455

Print the survey design summary

```
> dsgnsum(Stratsites)
```

Design Summary: Number of Sites

stratum	
Central Basin and Range	Colorado Plateaus
25	25
Mojave Basin and Range	Northern Basin and Range
10	10
Southern Rockies	Wasatch and Uinta Mountains
10	25
Wyoming Basin	Sum
10	115

6 Unstratified, unequal probability, GRTS survey design with an oversample and a panel structure for survey over time

The fourth survey design is an unstratified, unequal probability design with an oversample and a panel structure for survey over time. List `Paneldsgn` is assigned design specifications. Since the survey design is unstratified, `Paneldsgn` includes a single list named "None" that contains four items: `panel`, `seltype`, `caty.n`, and `over`. A vector identifying sample sizes for five panels is assigned to `panel`. The value "Unequal" is assigned `seltype`, which indicates unequal selection probabilities. The third item, `caty.n`, assigns sample sizes for each of seven multidensity categories, where ecoregion names

are used as the categories. Note that the sum of sample sizes provided in `caty.n` must equal the sum of sample sizes in `panel`. The value 100 is assigned to `over`, which specifies an oversample of 100 sites. An oversample is replacement sites for the survey design. The `grts` function attempts to distribute the oversample proportionately among sample sizes for the multidensity categories. If the oversample proportion for one or more categories is not a whole number, a warning message is printed and the proportion is rounded to the next higher integer. For this example, the oversample is not proportionate to the category sample sizes, and the warning message is printed by calling the `warnings` function.

For this survey design, a shapefile will be used as the sampling frame. The following arguments are included in the call to `grts`: (1) `design`: assigned the `Paneldsgn` list; (2) `DesignID`: assigned the value "UNEQUAL"; (3) `type.frame`: assigned the value "area"; (4) `src.frame`: assigned the value "shapefile"; (5) `in.shape`: assigned the value "eco_l3_ut"; (6) `att.frame`: assigned the `att` data frame; (7) `mdcaty`: assigned the value "level3_nam"; and (8) `shapefile`: assigned the value FALSE. Upon completion of the call to `grts`, the initial six sites for the survey design and a design summary are printed.

Create the design list

```
> Paneldsgn <- list(None=list(panel=c(Panel_1=50, Panel_2=50, Panel_3=50,
+                                     Panel_4=50, Panel_5=50),
+                                     seltype="Unequal",
+                                     caty.n=c("Central Basin and Range"=64,
+                                             "Colorado Plateaus"=63,
+                                             "Mojave Basin and Range"=15,
+                                             "Northern Basin and Range"=15,
+                                             "Southern Rockies"=15,
+                                             "Wasatch and Uinta Mountains"=63,
+                                             "Wyoming Basin"=15),
+                                     over=100))
```

Select the sample

```
> Panelsites <- grts(design=Paneldsgn,
+                    DesignID="UNEQUAL",
+                    type.frame="area",
+                    src.frame="shapefile",
+                    in.shape="reg1_lakes",
+                    att.frame=att,
+                    mdcaty="level3_nam",
+                    shapefile=FALSE)
```

Stratum: None

Initial number of levels: 5

Current number of levels: 5

Current number of levels: 7

Final number of levels: 7

Print the warning message

```
> warnings()
```

Warning message:

```
In grts(design = Paneldsgn, DesignID = "UNEQUAL", type.frame = "area", :
```

Oversample size is not proportional to category sample sizes for stratum "None".

Print the initial six lines of the survey design

```
> head(Panelsites@data)
```

	siteID	xcoord	ycoord	mdcaty	wgt	stratum
1	UNEQUAL-001	-1212767	2039497	Wasatch and Uinta Mountains	691099224	None
2	UNEQUAL-002	-1501559	2063485	Central Basin and Range	1282258537	None
3	UNEQUAL-003	-1147479	1672335	Colorado Plateaus	1361859732	None
4	UNEQUAL-004	-1435433	1784866	Wasatch and Uinta Mountains	691099224	None
5	UNEQUAL-005	-1287775	2102686	Wasatch and Uinta Mountains	691099224	None
6	UNEQUAL-006	-1427132	2240055	Northern Basin and Range	175999363	None

	panel	EvalStatus	EvalReason	level3	area_ha
1	Panel_1	NotEval		19	4256949
2	Panel_1	NotEval		13	8206455
3	Panel_1	NotEval		20	8579716
4	Panel_1	NotEval		19	4256949
5	Panel_1	NotEval		19	4256949
6	Panel_1	NotEval		80	263999

Print the survey design summary

```
> dsgnsum(Panelsites)
```

Design Summary: Number of Sites Classified by mdcaty (Multidensity Category) and panel

		panel					
mdcaty		OverSamp	Panel_1	Panel_2	Panel_3	Panel_4	Panel_5
Central Basin and Range		29	9	11	13	16	10
Colorado Plateaus		26	11	13	13	12	11
Mojave Basin and Range		6	3	3	3	4	4
Northern Basin and Range		7	3	2	5	3	3
Southern Rockies		6	1	3	3	4	4
Wasatch and Uinta Mountains		22	20	14	10	9	13
Wyoming Basin		6	3	4	3	2	5
Sum		102	50	50	50	50	50

		panel
mdcaty		Sum
Central Basin and Range		88
Colorado Plateaus		86
Mojave Basin and Range		23
Northern Basin and Range		23
Southern Rockies		21

Wasatch and Uinta Mountains	88
Wyoming Basin	23
Sum	352