METHODS FOR APPROXIMATING POPULATIONS

DESCRIPTION OF METHODS

1. *CENSUS:* Complete count of all individuals in a population - **Not applicable for this assignment – do not attempt**.

2. *SURVEY***:** A partial count; could be expanded to an estimate of total population size.

   a. *Mark/Recapture –* ***Refer to Krebs 1999, Chap. 2, for additional background (bCourses: Files > Field Project > Additional Resources)***
   Capture and mark a known number of individuals from a given population. Return the marked individuals to the population and allow a period of time for them to disperse back into the general population (depending on the species this could be hours or days). Take another sample and count the ratio of the marked to unmarked individuals to estimate your population.

   General Assumptions of Mark-Recapture studies
   - "Closed" population (i.e., no emigration/immigration/birth/death)
   - No loss of markers
   - Equal detectability of marked and unmarked individuals
   - Random mixing of marked individuals
   - Marked individuals are a representative sample of the population
   - All individuals are independent of one another

   NOTE: DUE TO CAMPUS ANIMAL CARE AND USE REGULATIONS, YOU MAY <u>NOT</u> CAPTURE AND MARK ANY VERTEBRATE SPECIES

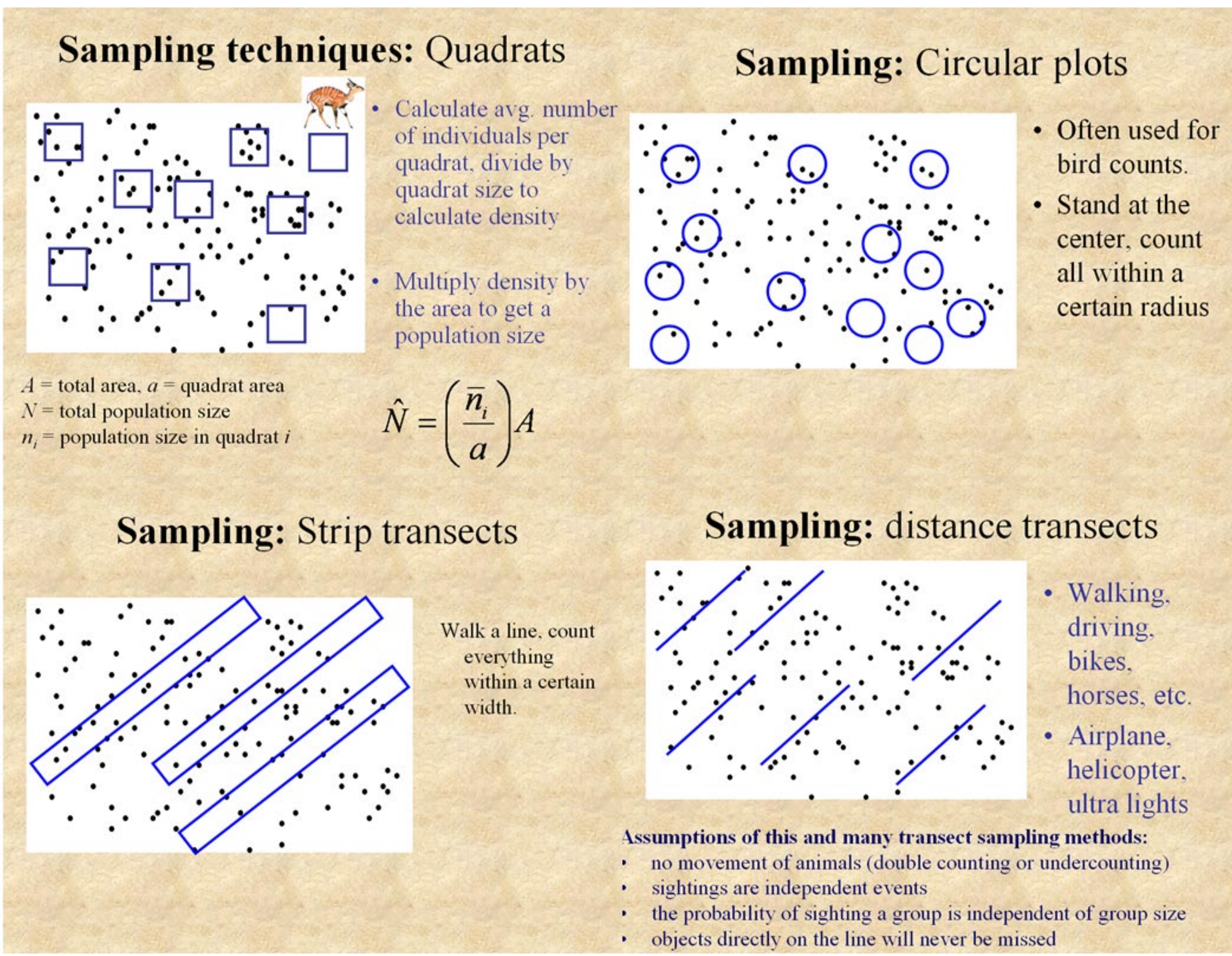
   b. *Plot/Transect sampling*
      a. *Quadrats:* Count the number of animals in randomly placed square quadrats. Works well for less mobile, smaller species (e.g., insects).
      b. *Circular Plots:* Count the number of animals within a certain radius of randomly placed circular plots. Commonly used for birds and referred to as "point counts."
      c. *Line/Strip Transects:* Travel along randomly placed lines and count animals seen within a given distance of the survey lines. Works best for medium-sized animals that are easy to see (e.g., raccoons at night).
      d. *Distance Transects:* Travel along randomly placed lines and record all animals seen *and* their distances from the line. This data can be used to calculate how detection probability varies with distance to provide an abundance estimate that is corrected for imperfect detection. **Not applicable for this assignment**, unless you are highly motivated and want to teach yourself how to use Program Distance (http://www.ruwpa.st-and.ac.uk/distance/).

   General Assumptions of Plot/Transect Sampling
   - Quadrats/circular plots/transects are representative of the entire area and are randomly placed. **NOTE**: the human mind is not random, so you actually need to use a method to make sure that sampling locations are randomized. For

example, you could obtain a map of your study area from Google Earth, divide your study area into numbered squares and randomly select the squares you will sample using a random number generator. The website http://random.org/ offers an easy way to generate a random number over any given interval.

- No movement of animals (double counting or undercounting)
- Sightings are independent events (sighting of one individual does not influence the sightings of others)
- All individuals are equally likely to be seen and the probability of sighting a group is independent of group size
- Objects directly on the line will never be missed (for distance sampling)

Note: For quadrats, circular plots, and strip transects, you probably will assume that all individuals within the searched area were detected, which may not be true. (More advanced methods use repeated surveys to estimate detection probability. You do NOT need to worry about estimating detection probability for this assignment.)



3. *INDEX (indirect sampling)*
   Choose an appropriate index that indirectly represents the population (pellet groups, road kills, scent stations, tracks, nests, etc.) and use the data to find trends in population (i.e. more dense populations in some areas compared to others, decreasing or increasing population over time). If you use this method, you must have some information, like a previous study, about how the index translates into numbers of individuals, and then make the conversion from numbers of indices to numbers of animals. For example, you may have evidence that there is an extremely high correlation between the number of nests and the number of feral pigs, and that each nest is associated with 4 pigs in a given area. Rather than counting each of the pigs (which may be difficult to detect), you can simply count the number of nests (the index), multiply by 4, and have a reasonably accurate estimate of the number of pigs in the sampled area with much less effort.

SAMPLE DATA SET #1: MARK-RECAPTURE

Let's say you want to estimate the number of snails in your yard. To do so, you capture 20 snails, mark them all, and release them back into the yard. Three days later, you conduct the second trapping of snails over an eight-hour period. On this trial, you catch 10 snails total, 3 of which are marked from the previous snail trapping. The table below summarizes the results obtained.

M	C	R	$\hat{N}$
20	10	3	56.75

M = total number captured and marked during first visit
C = total number captured during second visit
R = number of marked individuals seen on second visit

1. First, estimate the population size with the **unbiased Petersen estimator*** (*this is the recommended equation you should use in your study due to the small sample size*)

$$\hat{N} = \text{population estimate} = \frac{(M+1)(C+1)}{(R+1)} - 1 = 56.75 \text{ [or} \sim 57 \text{ snails]}$$

The estimator is unbiased if $(M + C) > N$ and nearly unbiased if there are at least seven recaptures of marked animals ($R > 7$). Since neither of these criteria is fulfilled in the dataset above, the estimate here is *probably* an overestimate of the actual population size. Thus, for your study it is recommended that you try to get at least 7 recaptures!

***Note:** This is an improvement over the simple Petersen estimator (MC/R) that leads to overestimates of population size.

2. Second, calculate the **confidence intervals** for your population estimate. For capture-recapture studies, confidence intervals can be calculated in various ways.
   a. Calculate the variance of the sample. For this assignment, we will use the approximately unbiased estimator of variance given by:

$$var(N) = \frac{(M+1)(C+1)(M-R)(C-R)}{(R+1)(R+1)(R+2)} = 343.6$$

   b. From this, we can calculate an approximate 95% confidence interval as:

$$95\% \text{ CI} = N \pm 1.96\sqrt{var(N)}$$

$$\text{Lower CI} = 56.75 - 1.96\sqrt{343.6} = 20.45 \text{ [or} \sim 20 \text{ snails]}$$

$$\text{Upper CI} = 56.75 + 1.96\sqrt{343.6} = 93.05 \text{ [or} \sim 93 \text{ snails]}$$

This assumes that the population size (N) is approximately normally distributed with large sample sizes. If you are confused about calculating confidence intervals, you can ask the GSIs for more help.

NOTE: In your calculations, you should carry the additional decimal places until the end, and then round to the nearest individual. In general, your estimate should not have more significant figures than the data that it is based upon.

Additional background information on the above statistics is available in the “**Additional Statistic Examples**” document (bCourses: Files > Field Project > Additional Resources)

SAMPLE DATA SET #2: QUADRAT ESTIMATION

Let's say that you want to estimate the number of banana slugs in Tilden Park. You set out 50 randomly located quadrats (x) and exhaustively search and count the number of slugs in each (n_i). You obtain the data shown here for the number of slugs in each plot:

3	1	0	2	0	2	0	2	0	0
2	1	0	1	0	0	0	2	1	1
0	0	0	0	0	1	0	1	1	0
0	0	4	0	1	0	0	0	0	1
2	0	0	1	0	1	0	0	0	0

First, calculate the **mean density** of slugs in your quadrats:

$$\text{Mean Density } = \frac{\sum n_i}{x} = \frac{31 \text{ slugs}}{50 \text{ quadrats}} = 0.62 \text{ slugs/quadrat}$$

Second, calculate the **standard deviation** of your data (how much variation there is in your data set relative to the sample mean)

$$\text{Standard Deviation (SD)} = \sqrt{\frac{\sum n_i{}^2 - \frac{(\sum n_i)^2}{x}}{x-1}} = \sqrt{\frac{61 - \frac{31^2}{50}}{49}}$$
$$= \sqrt{0.8527} = 0.92 \text{ slugs/quadrat}$$

Third, calculate the **standard error** of your data.

$$\text{Standard Error (SE)} = \frac{\text{SD}}{\sqrt{x}} = \frac{0.92}{\sqrt{50}} = 0.13 \text{ slugs/quadrat}$$

Lastly, calculate the **95% confidence intervals** for your estimated population density.

$$\text{Lower CI} = \text{mean } - \text{ (SE)} \times \text{(critical value)} = 0.62 - (0.13)(2.010) = 0.36 \text{ slugs/quadrat}$$

$$\text{Upper CI} = \text{mean} + \text{ (SE)} \times \text{(critical value)} = 0.62 - (0.13)(2.010) = 0.88 \text{ slugs/quadrat}$$

Determining the Critical Value for a data set

- The **critical value** (2.010) in the above confidence interval calculation comes from the t-distribution table provided below.
- The critical value **will vary** in your calculations depending on:
  - Confidence level (most commonly 95% confidence, and what you will use in your own calculations)
  - Degrees of freedom (sample size minus 1)

- In this case, we assume that there is a normal, or 2-tailed, distribution about the mean.
- For the example above:
  - Sample size was 50
  - Degrees of freedom: 49 (sample size minus 1)
  - To obtain our 95% CI, we need to look up the critical value from the t-distribution table below for degrees of freedom equal to 49 at the 95% confidence level
    - The table has values for degrees of freedom of 48 (2.011) and 50 (2.009), but not for 49, so we estimate the critical value between the two values given (2.010)

Convert density to population size and extrapolate to estimate the total population size of your study area

All of the calculations above were for the *density* of slugs. To get an estimate of the *total population size*, first convert your 'per quadrat' estimate into a unit of area then multiply the density by the total area. Suppose the total area was 10 km^2 and your quadrat size was 0.01km^2.

$$\text{Estimated population size} = \frac{0.62\ \text{slugs/quadrat}}{0.01\ \text{km}^2/\text{quadrat}} \times 10\ \text{km}^2 = 620\ \text{slugs}$$

$$\text{Lower confidence interval (CI)} = \frac{0.36\ \text{slugs/quadrat}}{0.01\ \text{km}^2/\text{quadrat}} \times 10\ \text{km}^2 = 360\ \text{slugs}$$

$$\text{Upper confidence interval (CI)} = \frac{0.88\ \text{slugs/quadrat}}{0.01\ \text{km}^2/\text{quadrat}} \times 10\ \text{km}^2 = 880\ \text{slugs}$$

Remember that for this to be a reasonable estimate, the quadrats must be "representative" of the entire area for which you are attempting to estimate the population size.

NOTE: In your calculations, you should carry the additional decimal places until the end, and then round to the nearest individual. In general, your estimate should not have more significant figures than the data that it is based upon.

Additional background information on the above statistics is available in the "**Additional Statistic Examples**" document (bCourses: Files > Field Project > Additional Resources)

T DISTRIBUTION TABLE

Confidence Level

Degrees of Freedom	80%	90%	95%	98%	99%	99.8%	99.9%	
1	3.078	6.314	12.710	31.820	63.660	318.300	637.000	1
2	1.886	2.920	4.303	6.965	9.925	22.330	31.600	2
3	1.638	2.353	3.182	4.541	5.841	10.210	12.920	3
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610	4
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869	5
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959	6
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041	8
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781	9
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587	10
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437	11
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318	12
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221	13
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140	14
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073	15
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015	16
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965	17
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922	18
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883	19
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850	20
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819	21
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792	22
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768	23
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745	24
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725	25
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707	26
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690	27
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674	28
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659	29
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646	30
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622	32
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601	34
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582	36
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566	38
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551	40
42	1.302	1.682	2.018	2.418	2.698	3.296	3.538	42
44	1.301	1.680	2.015	2.414	2.692	3.286	3.526	44
46	1.300	1.679	2.013	2.410	2.687	3.277	3.515	46
48	1.299	1.677	2.011	2.407	2.682	3.269	3.505	48
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496	50
55	1.297	1.673	2.004	2.396	2.668	3.245	3.476	55
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460	60
65	1.295	1.669	1.997	2.385	2.654	3.220	3.447	65
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435	70
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416	80
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390	100
150	1.287	1.655	1.976	2.351	2.609	3.145	3.357	150
200	1.286	1.653	1.972	2.345	2.601	3.131	3.340	200
Two Tails	0.20	0.10	0.05	0.02	0.01	0.002	0.001	
One Tail	0.10	0.05	0.025	0.01	0.005	0.001	0.0005	

Tail Probabilities