

## Basic Statistical Concepts



Kim Dietrich

## Introduction

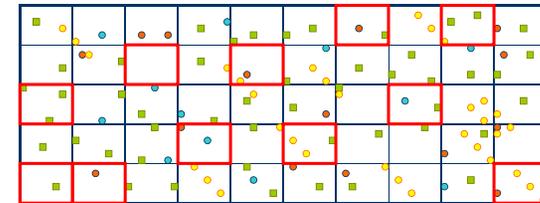
- Representative sample
  - Removes subjectivity / bias
  - Increased confidence in data
- Needed for estimating population size / assessing health
- Needed for managing fisheries

## Objectives

- Define random sampling & explain its importance
- List three levels where sampling occurs
- Describe how to use the random sample and random number tables
- Explain the difference between a random sample and a systematic random sample
- Demonstrate ability to choose a random sample and document sampling methodology

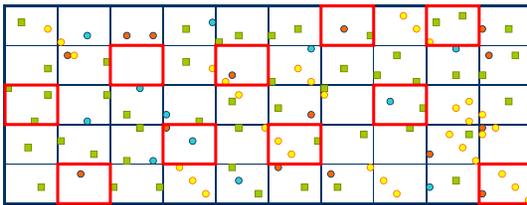
## What is a “statistically representative sample?”

- General – “selection of individual observations intended to yield some knowledge about a population of concern”
- Subset – used to make reliable predictions of population



## What is a “statistically representative sample?”

Symbol	Count in Sample	Extrapolated (sample * 5)	Actual
Green square	8	40	50
Red circle	5	25	15
Yellow circle	7	35	33
Blue circle	2	10	12



## What is a “statistically representative sample?”

- General – “selection of individual observations intended to yield some knowledge about a population of concern”
- Fisheries context – collect information on commercial catch in order to have knowledge about fish populations and health
- Random - every member of the population (catch) has an equal probability of occurring in the sample

## Sampling levels (strata)

- Fishery / gear type
  - Vessel
  - Trip
  - Haul or net
  - Species
- Agency
- Observer



## Sampling guidelines

- Collect before sorting
- Do not hand pick
- Collect from multiple points
- Larger sample better
- Selecting hauls – Random sample table (RST)

Hauls per day	RST	Target sample rate
1-2	None	100%
3-4	#1	70-75%
5+	#2	65%-70%

## Random sample table

- Choose one table per trip
- Complete for each haul

Random Sample Table #2

Date	Haul #	Samp?	Notes	Date	Haul #	Samp?	Notes
11-Nov-09	1	Y	watched left haul	12-Nov	9	Y	
11-Nov-09	2	Y		12-Nov	10	N	
11-Nov	3	Y		12-Nov	11	Y	
11-Nov	4	Y		13-Nov	12	Y	
11-Nov	5	N	Caught up on papers	13-Nov	13	Y	
12-Nov	6	N		13-Nov	14	Y	Sea Sick - weather
12-Nov	7	Y		13-Nov	15	N	
12-Nov	8	Y		13-Nov	16	N	

## Sampling Description

- Complete for each vessel and each change in sampling strategy
- Flow of fish
- Sample design
  - Selecting hauls
  - Within-haul sampling
  - Lengths/weights/age structures
  - Specimen samples

## Steps in Taking a Random Sample

1. Define population
2. Define sampling frame
  - Spatial – space or gear
  - Temporal
3. Define sample units
4. Number sample units
5. Decide how many units to sample
6. Randomly choose units (random numbers)

## Random Systematic Sampling

- Knowledge of total sampling units ideal
- Determine how many units you want to sample
- Divide total units by # units you want to sample
- Select a random number between 1 and result of above
- Sample every  $n^{\text{th}}$  unit thereafter

## Random number table

- Appendix 11 in manual
- Enter at random point
- Determine # digits
- Determine direction
- Example

4	6	2	3	9	3	2	4	3	5	1	4	7	2	9
9	7	6	9	7	6	4	1	9	9	2	1	9	0	8
8	0	5	4	3	0	8	6	3	2	7	6	1	1	6
4	7	4	6	6	9	3	2	1	6	6	0	6	3	8
1	8	6	2	1	3	7	7	1	5	3	5	4	4	1
0	8	2	8	4	6	3	8	9	2	9	0	7	0	2
4	0	5	4	5	0	0	7	9	9	1	1	1	6	4
4	8	8	8	8	7	3	4	8	8	8	7	5	3	6
8	4	8	6	9	2	6	2	6	7	3	7	0	9	8
8	0	8	1	3	8	4	3	4	7	4	1	7	3	2
8	3	9	5	9	7	5	8	2	2	3	9	8	8	8
6	3	2	6	0	1	7	5	6	6	6	6	2	3	
3	1	0	9	0	8	9	3	0	5	5	6	5	7	6
3	2	5	4	9	9	0	8	3	9	4	4	1	6	7

## Random sample - example

- 30 units – number in advance
- Target sample rate 20%
  - How many units? **6**
- Select 6 random numbers between 1 and 30

1	3	4	4	1	3	7	5	1	1	9	1	3	4	7
0	1	4	1	4	4	5	1	0	4	8	9	6	0	8
0	2	5	8	9	6	6	5	5	5	4	0	6	2	7
1	7	5	1	3	4	4	0	3	5	9	0	9	9	5
9	8	4	4	2	3	5	7	1	6	4	9	5	0	8
2	9	8	8	8	9	6	8	1	3	6	0	9	1	6
8	0	1	0	5	6	7	1	2	9	3	0	0	8	1
7	8	8	4	4	4	1	7	8	1	1	9	1		
1	0	0	1	7	2	2	9	8	1	1	5	3	4	7
7	2	1	9	4	2	8	3	8	1	8	0	8	4	0



## Random systematic sample - example

- 30 units – number in advance
- Target sample rate 20%
  - **30 units / 6 units = 5**
- Select random number between 1 and 5 > **1**
- Sample 1<sup>st</sup> unit & every 5<sup>th</sup> unit thereafter
  - Units **1, 6 (1+5), 11 (6+5), 16 (11+5), 21 (16+5), 26 (21+5)**



## Sample bias

- Catch stratification
  - In nets or bins
  - Across depth strata
- Crew sorting
- Collection location / mechanical bias – selection by hand, size of shovels, incline belts/doors
- Sample size



## Activity #1

- Work in groups of 2
- Label units on handout 1 to 100 (this can be a mental label)
- Create a sampling plan based on the sample rate and type being handed out
- Circle the quadrants you sample
- Answer the questions on handout

## Activity #2

- To be determined

## Summary

- What is random sampling
- Why is random sampling important?
- List three levels where sampling occurs
- Describe how to use the random sample table
- When should you use a random number table?
- What is the difference between a random sample and a systematic random sample?

## Activity – Sampling symbols using quadrants

Names:

### Instructions:

1. Label quadrant units 1 to 100
2. Create a sampling plan based on the sample rate (10%, 25%, 50%) and type (random or systematic random) being handed out
3. Select the “quadrants” to sample using the random number table below

7 0 7 4 4	7 0 6 6 9	6 5 2 5 1	9 9 1 2 8	7 5 0 6 1
7 7 3 3 4	1 0 3 3 6	6 3 8 7 5	0 6 5 7 5	7 7 9 8 2
0 4 6 2 7	3 6 7 4 4	7 1 4 7 2	2 2 9 8 8	0 5 7 0 7
4 2 9 8 7	9 9 1 3 4	7 8 8 2 1	6 5 1 7 1	0 9 6 0 1
3 1 9 4 3	3 2 0 9 0	1 0 2 7 9	0 4 2 6 2	7 9 5 9 5
3 5 1 9 4	0 4 8 9 4	6 9 7 3 0	4 0 8 6 9	9 2 3 4 3
6 9 8 6 6	4 0 0 1 2	2 5 3 3 5	7 1 2 3 7	3 1 0 2 6
6 4 9 3 2	3 9 4 0 4	8 2 5 5 2	6 1 3 7 2	4 7 6 6 7
0 3 5 7 1	4 2 2 2 4	0 4 9 7 2	4 8 1 5 5	7 6 8 7 6
6 4 0 2 7	7 0 9 0 1	5 7 4 4 8	4 1 3 2 6	8 2 7 9 6
7 0 8 3 9	4 5 1 5 0	9 5 6 6 3	7 8 5 4 2	4 8 5 2 4
4 0 4 4 6	6 3 9 3 3	0 9 8 1 3	9 7 1 8 3	0 7 6 6 6
2 1 8 9 6	7 7 0 7 8	5 9 3 9 7	1 9 3 4 8	5 9 8 9 0
9 9 6 0 9	7 6 3 0 4	3 3 2 5 0	1 6 1 9 2	2 6 0 4 6
9 1 1 9 3	8 4 7 9 6	7 1 0 4 7	8 5 3 6 7	4 3 5 1 9

4. Circle the quadrants you sampled
5. Answer the questions below

### Questions:

1. What is your population?
2. What is your sample frame and unit?
3. What is your target sample rate?

4. If random sample, what were your random numbers?
5. If systematic sample, what was your random number and which units did you sample?
6. Count each of the shape types in your sample & record below:

Shape	Tally	Sample total
Lightning bolt 		
Clubs 		
4-point star 		
crescent 		
Sun 		

7. What extrapolation factor do you need to multiply by to calculate an estimate of the number of each shape in the population? \_\_\_\_\_
8. Extrapolate counts to total population

Shape	Estimate of population total
Lightning bolt	
Clubs	
4-point star	
crescent	
sun	

# Sampling - Activity #1

