Areas of Tails of Distributions

Definition

The left tail of a density curve $y = f(x)$ of a continuous random variable X cut off by a value x^* of X is the region under the curve that is to the left of x^*. The right tail cut off by x^* is defined similarly

The probabilities tabulated in the cumulative probability table for Z are areas of left tails in standard normal distribution

Tails of the Standard Normal Distribution:

At times, it is important to solve problems where a specific area is known, and we need to determine the corresponding value z^* of Z. This requires reading the cumulative probability table in reverse, locating the relevant area, and identifying the corresponding Z value.

Example

find z^* such that $P(Z > z^*) = 0.0250$

solution; Since this is a right tail, we compute $1 - 0.0250$: 0.9750 and locate 0.9750 in the table. It corresponds to $z^* = 1.96$

find $z_{.01}$ and $-z_{.01}$ for a right and left tail area of 0.01

Solution: Since $-z_{.01}$ cuts off a left tail, we look for 0.0100 in the table. It lies between 0.0102 and 0.0099 with $z^* = -2.33$. By symmetry, $z_{.01} = 2.33$

Tails of General Normal Distribution

for a normally distributed X with mean μ and standard deviation σ, the value x^* cutting off a left or right tail of area c is found as follows

1) find z^* that cuts off the area c in the standard normal distribution
2) Use the formula $x^* = \mu + z^* \sigma$

Example
find x^* such that $P(X > x^*) = 0.65$ with $\mu = 175$, $\sigma = 12$

solution; compute $1 - 0.65 = 0.35$ searching for 0.3500 in the table gives $z^* = -0.3950$

$x^* = 175 + (-0.39) * 12 = 170.32$

Example find the minimum GRE score for the top 5% of scores, given $\mu = 510$, $\sigma = 60$

solution compute $1 - 0.0500 = 0.9500$, searching the table gives $z^* = 1.645$

so $x^* = 510 + 1.645 * 60 = 608.7$

Summary
To determine a value x^* or z^* that corresponds to a given probability in a normal distribution

1) Standard normal distribution
- To find z^* when $P(Z < z^*) = C$ locate C in the
- cumulative probability table and read corresponding z^* for right tail probabilities, use $1 - C$ to find the corresponding z^*

2) General Normal distribution
- Convert the problem into a standard normal form by finding z^* first
- Use the transformation formula $x^* = \mu + z^* \sigma$ to find x^*

3) Examples
- Calculating cutoffs for left and right tails using cumulative probability tables.
- Applying normal distribution principles to real-world scenarios like test scores and race times

Date / /
Page No.

Introduction to Sampling Distributions

1) Inferential Statistics and Sampling Distributions

- Inferential statistics involve generalizing from a sam to a population
- Example: Suppose you randomly sample 10 women aged 21-35 in Houston and calculate the mean height.
  - The sample mean will be different from the population mean
  - A second sample of 10 women will also likely yield a different mean
- The goal is to estimate how much sample statistics vary from one another and from the true population parameter
- This variation is analyzed using sampling distribution

2) Discrete Sampling Distribution

Example Using Pool Balls:

Three pool balls are numbered 1, 2 and 3

Two balls are randomly sampled with replacement, and their mean is computed

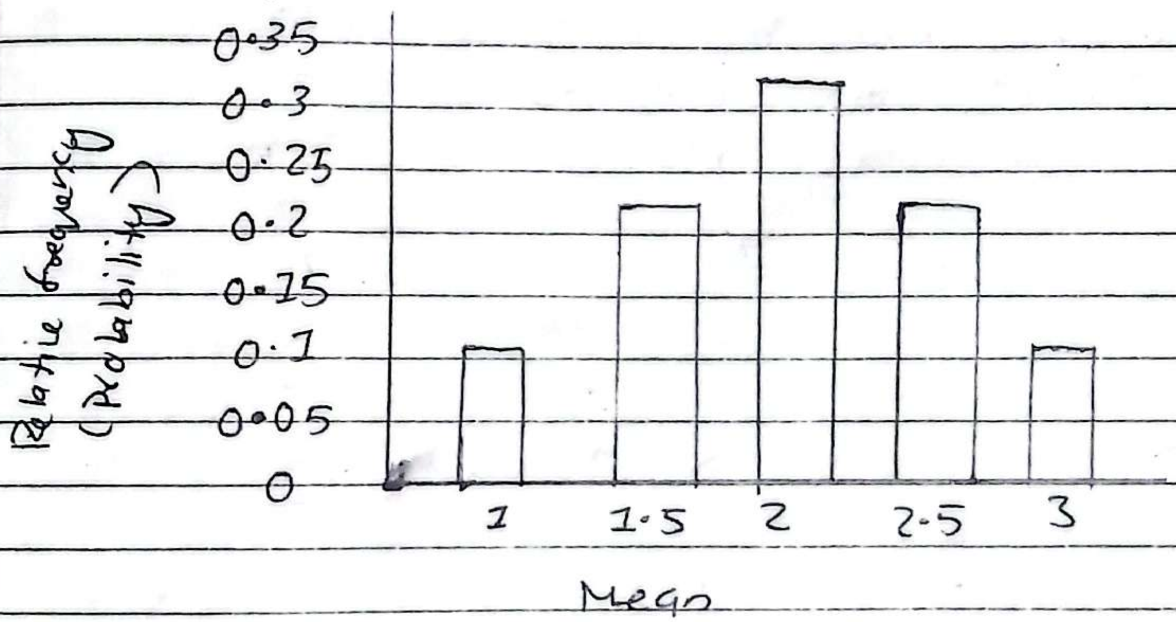
Possible outcomes

- The possible sample means are 1.0, 1.5, 2.0, 2.5 and 3.0

Frequency distribution (Table 2):

- The relative frequency dist of each mean is determ based on all possible outcomes (9 in total)

Mean	Frequency	Relative Frequency
1.0	1	0.111
1.5	2	0.222
2.0	3	0.333
2.5	2	0.222
3.0	1	0.111



Graphical Representation
The relative frequency distribution of sample means forms the Sampling distribution of the mean. This is also a probability distribution, where the y-axis represents the probability of obtaining each mean.

3) Conceptualizing Sampling distribution
Alternative perspective:
- Imagine repeatedly drawing two-ball samples and computing the mean for thousands of samples
- As the number of samples increases, the relative frequency distribution approaches the true sampling distribution

- It the number of samples approaches infinity, the relative frequency distribution becomes the exact sampling distribution

4) Sampling Distribution of other Statistics (Range Example)

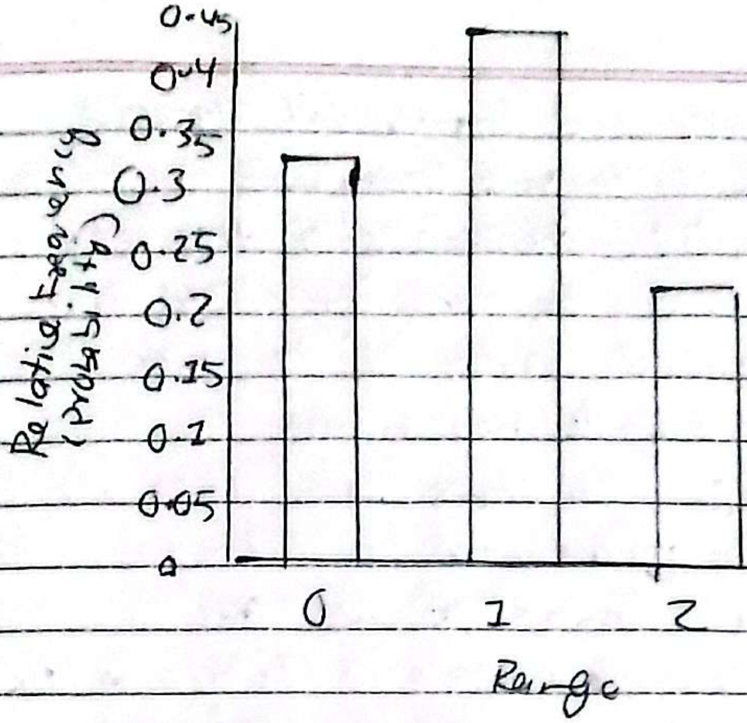
- Every statistic has a sampling distribution, not just the mean
- Example: Range of two sampled Balls
  Possible ranges and outcomes

Outcome	Ball 1	Ball 2	Range
1	1	1	0
2	1	2	1
3	1	3	2
4	2	1	1
5	2	2	0
6	2	3	1
7	3	1	2
8	3	2	1
9	3	3	0

The Range is computed as the difference between the larger and smaller sampled number. Possible range values 0, 1 or 2

Frequencies of Ranges for N = 2

Range	Frequency	Relative Frequency
0	3	0.333
1	4	0.444
2	2	0.222



Graphical Representation
Similar to means, the sampling distribution of the ~~range~~
can be graphed
Sampling distributions exist for different sample sizes
(eg, N=3)

5) Continuous Sampling Distribution
Example : Instead of just three pool balls, consider 1000
pool balls with values between 0.001 and 1.000.
There are now 1000 X 1000 = 1,000,000 possible
sample outcomes

Key Differences from Discrete Distributions
- listing all possible outcomes becomes impractical
- Instead, sampling distributions are estimated using relative frequency distribution from repeated samples
- Continuous distributions use probability densities instead of discrete probabilities

Date / /
Page No.

6) Sampling Distributions and Inferential statistics
In practice, sampling works in reverse:
1) Collect a sample and compute statistics (eg mean)
2) Use this to estimate the sampling distribution
3) Estimate the standard error of the mean (SEM), which measures how much sample means vary.
4) Use this knowledge to infer how close the sample mean is to the population mean
Example: Standard Error of the Mean
- Suppose your sample mean is 125 and the SEM is 5
- If the distribution is normal, the true population mean is likely within 10 units of the sample mean (since most values fall within two standard deviations

7) Beyond the Mean: Other Sampling Distributions
All statistics have sampling distribution including Variance, Difference between means, Pearson's Correlation Coefficient
These distributions help in estimating population param[eters] and conducting hypothesis tests.