AP Statistics – Unit 1 Quick Notes

Exploring One-Variable Data

1. Types of Variables

Categorical (Qualitative)

- Nominal – no order (e.g., species)
- Ordinal – order but not equally spaced

Quantitative (Numerical)

- Discrete vs Continuous
- Ratio scale – true zero, ratios meaningful
- Interval scale – arbitrary zero, ratios meaningless

2. Graphical Representations

Categorical: frequency table, relative freq., bar chart, pie chart. **Quantitative:** dotplot, stemplot, histogram, boxplot, CDF plot, density plot.

Histogram vs Bar Chart:

- Histogram – quantitative, contiguous bins
- Bar chart – categorical, separated bars

3. Describing Distributions – SOCS

- **Shape:** symmetric, skewed left/right, unimodal, bimodal, uniform
- **Outliers:** gaps, clusters, extreme values
- **Center:** median (resistant), mean (non-resistant)
- **Spread:** range, IQR, SD

Always describe **in context** and reference a visual.

4. Measures of Center & Spread

Center

$\bar{x} = \frac{1}{n}\sum x_i$ (sample mean)
Median = 50th percentile
Mode = most frequent

Spread

$s^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ (sample variance)
$s = \sqrt{s^2}$ (sample SD)
IQR $= Q_3 - Q_1$
Range = max − min

Outlier Rules: $x < Q_1 - 1.5 \times IQR$ or $x > Q_3 + 1.5 \times IQR$ $|z| > 2$ can also indicate unusual values.

5. Comparing Distributions

For each group: use SOCS *and* explicitly compare shape, center, spread, outliers.

6. Normal Distributions

Empirical Rule

68% within 1σ, 95% within 2σ, 99.7% within 3σ.

Chebyshev's Theorem

At least $1 - \frac{1}{k^2}$ within $k\sigma$ (any shape).

Z-score: $z = \frac{x-\mu}{\sigma}$ $\quad (z > 0$ above mean, $z < 0$ below mean)

Percentiles $\leftrightarrow$ z-scores: Can solve for μ or σ from known percentiles.

7. Effects of Transformations

- Add c: mean $\to +c$, spread unchanged
- Multiply by a: mean $\to \times a$, variance $\to \times a^2$, SD $\to \times |a|$
- Scaling changes fences and values proportionally $\Rightarrow$ same # of outliers

8. Special Tools

Normal Probability Plot: roughly straight line $\Rightarrow$ approx.

9. Useful Calculator Commands (TI-84 / TI-Inspire)

Data Entry:

- `STAT` $\to$ `EDIT` – enter data in a list

Summary Statistics:

- `STAT` $\to$ `CALC` $\to$ `1-Var Stats` – gives $\bar{x}$, s, Q_1, Q_3, min, max

Graphs:

- `2nd Y=` (`STAT PLOT`) – turn on plot
- Choose: boxplot (modified shows outliers), histogram, scatterplot

Outlier Detection:

- Use modified boxplot ($*$ marks outliers)

Normal Distribution:

- `2nd VARS` (`DISTR`)
- `normalcdf(lower, upper,` μ`,` σ`)` – area under curve
- `invNorm(area,` μ`,` σ`)` – percentile $\to$ value

Z-scores:

- Compute manually: $(x - \mu)/\sigma$ or store in list