STAT1000 · Midterm 2 cheat sheet

From the M&M Stats Guide. Colors match the chapters.

Random variable?

- Fixed, already known → **not** a random variable
- Count it → **discrete**
- Measure it (time, distance, weight, temperature) → **continuous**

Probability table

- $\Sigma P(x) = 1$ missing one $= 1 -$ rest
- $\mu = \Sigma\, x \cdot P(x)$
- $\text{Var} = \Sigma(x-\mu)^2 P(x) = \Sigma x^2 P(x) - \mu^2$
- $\sigma = \sqrt{\text{Var}}$ (don't forget the root)
- at most ≤ · at least ≥ · more than > · fewer than <

Shift and stretch: $Y = a + bX$

- $\mu_Y = a + b\mu_X$
- $\sigma_Y = |b|\sigma_X$ $\text{Var}(Y) = b^2\text{Var}(X)$
- °C = (°F − 32)·5/9 → mean: plug in; SD: × 5/9 only

Binomial (BINS)

- Binary · Independent · *n* fixed · Same *p*
- Write: "*X* = number of … out of *n*" and "*X* ~ Bin(*n*, *p*)"
- $P(X = k) = \text{C}(n,k)\, p^k (1-p)^{n-k}$
- $\mu = np$ $\sigma = \sqrt{np(1-p)}$
- $P(X \geq 1) = 1 - (1-p)^n$
- $P(A \mid B) = P(A \text{ and } B) \,/\, P(B)$

68–95–99.7

- Pieces from the middle out, each side: **34 · 13.5 · 2.35 · 0.15**
- Percentiles at −3σ…+3σ: 0.15 · 2.5 · 16 · 50 · 84 · 97.5 · 99.85
- Uneven interval → add pieces. "How many" → % × total

z-scores

- $z = (x-\mu)/\sigma \;\Leftrightarrow\; x = \mu + z\sigma$
- $\sigma = (x-\mu)/z \;\Leftrightarrow\; \mu = x - z\sigma$
- Table = area **left**. Right = 1 − table. Between = table(*b*) − table(*a*)
- Top *p*% → look up 1 − *p*. Bottom *p*% → look up *p*
- Top 10% 1.28 · 5% 1.645 · 2.5% 1.96 · 1% 2.33
- Write: $P(X > a) = P(Z > (a-\mu)/\sigma) = \ldots$

Sample mean $\bar{x}$

- $\text{mean} = \mu$ $\text{SE} = \sigma/\sqrt{n}$ $(n \times 4 \rightarrow \text{SE} \div 2)$
- Shape: normal population → normal, any *n*. Otherwise *n* ≥ 30 → approx. normal (CLT). Else: can't say
- $z = (\bar{x} - \mu) \,/\, (\sigma/\sqrt{n})$
- Spread of $\bar{x}$ is always less than the population's

Words

- **P**arameter ↔ **P**opulation (*μ*, *σ*, *p*): fixed
- **S**tatistic ↔ **S**ample ($\bar{x}$): changes every sample
- Sampling distribution = distribution of a **statistic**

Normal problem, every time

1. Write the model: $X \sim \text{N}(\mu, \sigma)$, or for a sample mean $\bar{X} \sim \text{N}(\mu, \sigma/\sqrt{n})$
2. Sketch the bell, shade the area asked for
3. z to 2 decimals, write $P(Z < \ldots)$, use the table
4. Left = table · Right = 1 − table · Between = subtract
5. Backwards: area left → inside table → $z \rightarrow x = \mu + z\sigma$
6. Answer with units; check the side of the mean

Read the words

- "at least one" → 1 − P(none)
- "given that" → divide by P(given)
- "sample mean" / "average of n" → σ/√n
- "how many" → % × total
- "top p%" → look up 1 − p
- Binomial: define X in words + name Bin(n, p)