The last test (12/12 in class)

Continuous distributions

- Should know the definition of probability density and cumulative distribution
- Compute with uniform distribution
- Exponential definition of hazard rate.
- Normal (Gaussian) distribution - The formula for probability density, mean, and variance

Exercises

Chapter 5 (shipped and covered)
K^2 and t-tests will provide formulas for more examples of statistics (next page).

Should know the statistics story to the extent covered (if there is a question about statistics, will provide tables). Candy = t with 1 degree of freedom.

Jointly distributed random variables - discrete, jointly continuous

Joint probability density, marginal densities

Independent jointly distributed random variables

Transformations of jointly distributed random variables (in particular arithmetic operations on random variables)
Multinormal distribution
Multivariate Gaussian (no main axes) \textit{Chapter 7}

- Expectation is additive
- Variance is additive for independent random variables
- Covariance

- What the strong law of large numbers and the central limit theorem say \textit{Chapter 9}

The central limit theorem
Let \( X_1, X_2, X_3, \ldots \) be a sequence of identically distributed independent random variables, each having expected value \( \mu \) and variance \( \sigma^2 \).

Then

\[
\lim_{m \to \infty} P \left\{ \frac{\sqrt{m}}{\sigma} \left( \frac{X_1 + \ldots + X_m}{m} - \mu \right) \leq a \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} \, dx.
\]

---

**Example** (the statistical Z test):

Measuring the distance to a star (light-years)

We want to make a sense of measurement and take the average.
Outcomes (values) of measurement are independent identically distributed random variables, common mean is $\mu$ (unknown) and common variance $\sigma^2 = 4$ (light-years$^2$). How many measurements do we have to make to be reasonably sure \(95\%\) we are within $\pm 0.5$ light-years$^2$?

\[ \text{Solution: } Z_m = \frac{\sum_{i=1}^{N} X_i - \mu \cdot \sigma^2}{2 \sqrt{m}} \rightarrow \text{standard Gaussian.} \]
\[ p \left\{ -0.5 \leq \frac{\sum_{i=1}^{n} X_i}{\sqrt{n}} < 0.5 \right\} = 1 - \Phi \left( \frac{\sqrt{n}}{4} \right) \]

\[ p \left\{ -0.5 \frac{\sqrt{n}}{2} \leq Z_n \leq 0.5 \frac{\sqrt{n}}{2} \right\} \approx \Phi \left( \frac{\sqrt{n}}{4} \right) - \Phi \left( -\frac{\sqrt{n}}{4} \right) \]

\[ \Phi \left( a \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx = 2 \Phi \left( \frac{\sqrt{n}}{4} \right) - 1 \]

Cumulative distribution of the standard Gaussian

\[ 2 \Phi \left( \frac{\sqrt{n}}{4} \right) - 1 = 0.95 \]

\[ \Phi \left( \frac{\sqrt{n}}{4} \right) = 0.975 \]

Table: \( \frac{\sqrt{n}}{4} = 1.96 \)
\[ n = (7.84)^2 \approx 61.47 \]

Answer: 62.