• **Time and Places of Exam:** The exam will be in class, on Thursday, Feb 13 2003.

• **Exam Materials:** The exam is closed book. A few pages of formulas and tables will be appended to the exam.

  You can also prepare your own formula sheet: a single one-sided 8x11 sheet on which you can write whatever you like. You should bring a calculator.

• **Exam Coverage:** Questions for the exam will be based on material in the following sections of your textbook:

  Chapter 2 : Entire chapter.

  Chapter 3 : Sections 2-8.

  Chapter 4 : Sections 1-10.

  Chapter 5 : Sections 1-9.
Descriptive Statistics, Chap 2

• Types of data: qualitative and quantitative

• population and sample, parameter and statistics

• mean, variance and standard deviation of a sample

• median, lower (upper) quartile and 100pth percentile

• z-score, IQR and outliers

• Boxplot, histogram and bar plot
Boxplot

- Calculate the median $m$, lower and upper quartile, $Q_l$ and $Q_u$, and the interquartile range, IQR.

- Draw a box with ends at $Q_l$ and $Q_u$ and draw a vertical line inside the box to locate the median $m$.

- Draw whiskers out to $Q_u + 1.5\text{IQR}$ and $Q_l - 1.5\text{IQR}$, but trim them back to the most extreme data point in those ranges.

- Draw dots for each individual data point outside the box and whiskers.

Example: Ex 2.54
Probability, Chap 3

• Sample space, events, Compound events (union, intersection and complement)

• Probability, conditional probability and probability rule for compound events

• Independence and mutually exclusive

• Bayes theorem

• Counting rules
Example: In the week following final exams, 10% of Duke students go to the beach (it always seems like more); 50% go home; and 40% start a job. Of the beach-goers, 90% (claim to) have fun; 22% of the home-goers have fun; and a disappointing 25% of the workers have fun.

• a. What is the probability that a randomly-selected Duke student will go to the beach and not have fun?

• b. What is the probability that a randomly-selected Duke student will have fun (anywhere)?

• c. What is the probability that a randomly-selected Duke student who is having fun can see the ocean?

• d. Don’t you wish you had a job working at home and that you lived at the beach?
Example: A communications satellite is launched into orbit on an three-stage rocket. The first stage of the rocket fails with probability 0.10 and the second and third stage are more reliable, each failing with probability 0.01. The satellite itself has a rocket, which places it into geosynchronous orbit, failing with probability 0.001. Assuming that all stages of the rocket, and the satellite itself, fail (and succeed) independently:

- a. What is the probability that the satellite is successfully placed into the desired orbit? (0.8812)

- b. Given that the first stage works successfully, what is this success probability? (0.9791)

- c. Given that the first stage does not fail, what is the probability that the second stage does not fail? (0.99)

- d. You are told that the satellite did not make it into orbit, but nothing more. What is the probability that the first stage failed? (0.8418)
Counting Rule

• Basic rule: Multiplicative rule

• Important setting:
  – With replacement or Without
  – Order matters or not

• Other rules:
  – Permutations $n!$
  – Combinations (choose $k$ out of $n$) $\frac{n!}{k!(n-k)!}$.
  – Partitions $\frac{n!}{n_1! \cdots n_k!}$
Ex 3.44: The Federal Aviation Administration (FAA) has a 16-member committee working on evaluating the traffic control systems of four facilities replying on computer-based equipment. How many possible ways the FAA can form the task force to do the evaluation,

- a. if the FAA wants to assign one member to each facility and assume that each member can be assigned to more than one facilities?

- b. if the FAA wants to assign one member to each facility and assume that each member can NOT be assigned to more than one facilities?

- c. if the FAA wants to form a 4-memeber task force for all the four facilities?

- d. if the FAA wants to assign 4-member task force to each facilities?

a. $16^4$

b. $P_{4}^{16} = (16)(15)(14)(13)$

c. $(\begin{pmatrix} 16 \\ 4 \end{pmatrix}) = \frac{(16)(15)(14)(13)}{(4)(3)(2)(1)}$

d. $\frac{16!}{4!4!4!4!}$
Random Variables, Chap 4-5

- Discrete or continuous

- Probability distribution: probability mass function (discrete), probability density (continuous)

- Cumulative distribution function and its use in calculating probabilities for continuous random variables

- Expectation of functions of random variables: mean, variance (standard deviation)
Important Distributions

Discrete

- Bernoulli($p$), $x = 0, 1$
- Binomial($n, p$), $x = 0, 1, \ldots, n$
- Geometric($p$), $x = 0, 1, \ldots, n, \ldots$
- Poisson($\lambda$), $x = 0, 1, \ldots, n, \ldots$
- Discrete uniform

Continuous

- Uniform($a, b$), $x \in [a, b]$
- Normal($\mu, \sigma^2$), $x \in \mathbb{R}$
- Exponential($\beta$), $x \geq 0$
Distributions and Stories

Semiconductor Wafer: A semiconductor wafer contains a contamination particle with probability $p = 0.01$.

- Whether a new produced wafer is contaminated or not – Bernoulli($p$)

- How many are contaminated among 1000 wafers? – Binomial($n, p$)

- How many wafers will we look at before we find a contaminated one? – Geometric($p$)

- How many wafers will we look at before we find the $k$th contaminated wafer? – Negative Binomial

- Suppose today’s production is 1000 wafers and among them 10 are contaminated. Now randomly draw a sample of 100 wafers from today’s production, how many are contaminated? – Hypergeometric
**Fishing Example:** You go to a spot in Duke Forest to try to catch some fish. From past experience you know that the average number of fish you catch per hour is \( \lambda = 6 \).

- How many fish you catch in one hour of fishing? – Poisson(\( \lambda \))

- How many fish you catch in 20 minutes? – Poisson(\( \lambda/3 \))

- How long you have to wait to catch your first fish? – Exponential(\( \beta \)) with \( \beta = 1/\lambda \).

- How long you have to wait to catch your \( \alpha \)th fish? – Gamma(\( \alpha, \beta \)) with \( \beta = 1/\lambda \).
Normal Distribution

- Two parameter: mean $\mu$ and variance $\sigma^2$ (standard deviation $\sigma$)

- If $x \sim N(\mu, \sigma^2)$ then $(x - \mu)/\sigma$ is distributed as the standard normal $N(0, 1)$

- Calculation of normal probabilities using the standard normal random variable and the table in Appendix
Example: The number of cracks in a section of $I - 85$ which need to be repaired is assumed to be a Poisson random variable with parameter $\lambda = 1$ crack per mile.

- a. What’s the probability that no cracks require repair in 5 miles of highway? (0.0067)

- b. What’s the probability that more than one crack requires repair in 5 miles of highway? (0.9596)

- c. The highway is inspected and repaired in one mile section. What’s the probability that no cracks are found until mile 5? (0.0116)

- d. What’s the probability that 1 of the first 5 sections inspected has cracks, but the other 4 do not? (0.0579)
Example: The life of a semiconductor laser is normally distributed with a mean of 7000 hours and a standard deviation of 600 hours

• a. What’s the probability that a laser lasts less than 6000 hours? (0.0475)

• b. What’s the probability that a laser lasts more than 8000 hours? (0.0475)

• c. What’s the lifetime in hours exceeded by 99% of all lasers? (5602)

• d. Seven lasers are required in an industrial system. What’s the probability that all seven function for at least 6000 hours? ((1 – 0.0475)^7)