Design IV

E232 Fall 07

Class 17

Bruce McNair

bmcnair@stevens.edu

Notes on Midterm Grades

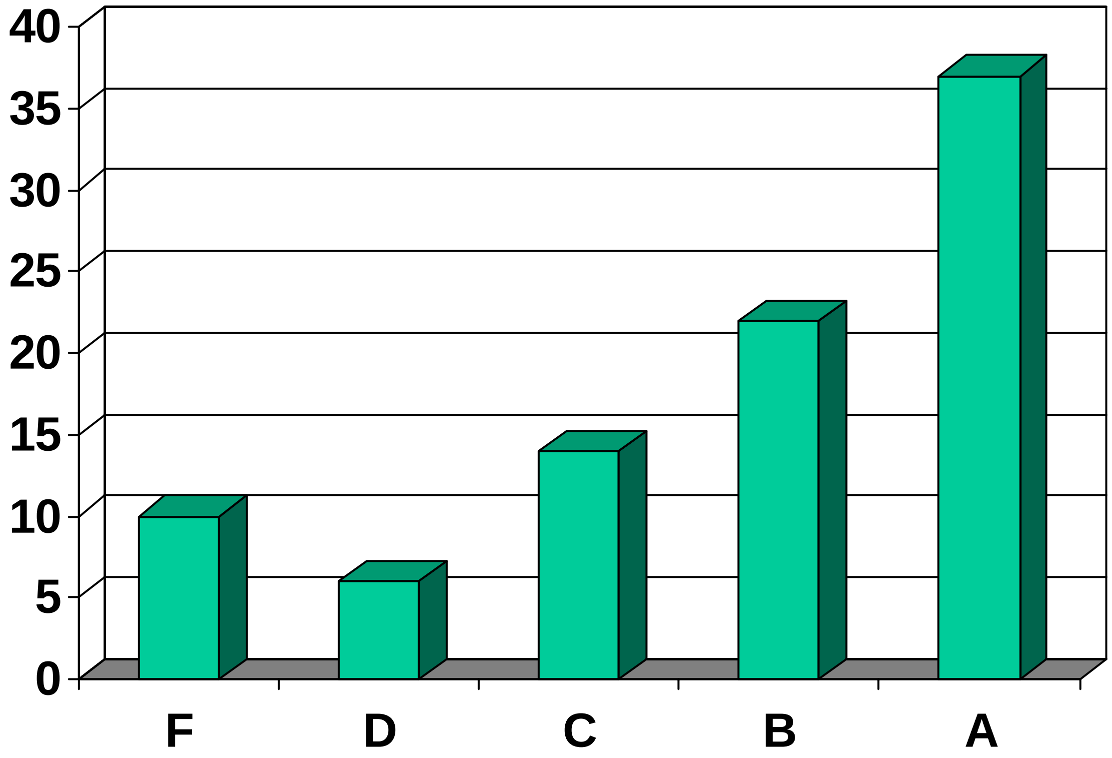
- All 100 and 200 level course instructors are required to assign student midterm grades
- Midterm grades are used to alert students to areas they need to address, they are not on your final transcript
- E232 midterm grades were whole letter grades only, no + or – grades were assigned. The standard letter grade levels were used without adjustment.
- For the E232 midterm grade, lab grades weren't included
- The lecture grade was based on quiz plus number of homework assignments submitted (and pledged), weighted approximately according to final value in course.

Notes on Midterm Grades

- In E232, the final grade will be 2/3rds lecture and 1/3rd lab grade (as noted in Class 1)

- I always grade my courses in the same manner (for final grades):
  - Calculate the overall score
  - Lookup the standard letter grade corresponding to that score
  - Examine the letter grade distribution and boundaries between letter grades
  - Adjust the mapping of numeric grades to letter grades to fit actual population data and special cases, adjusting the mapping downwards only, (e.g. while the standard distribution requires a 95 to get A, I might decide that 93.5 is sufficient, but will never increase the level above 95)
  - The final assignment of numeric scores to letter grades will ***always*** be monotonic.

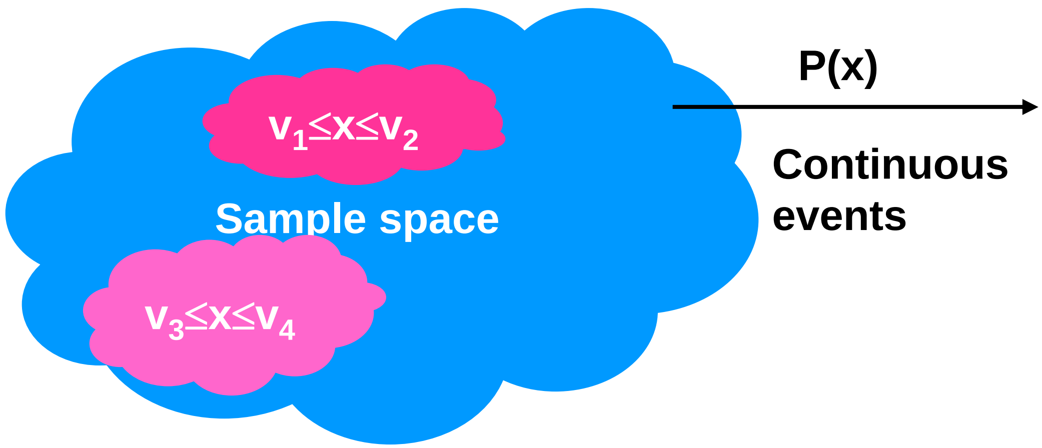
Notes on Midterm Grades



Regarding less than desired grades:

- Two students did not turn in Q1
- Nine students have not turned in the majority of weekly assignments

Probability Density Function



Experiment: Temperature measured by sensor

$$P(-\infty \le x \le \infty) = 1$$

- **x must have <u>some</u> value**

$$\mu = \sum_{i=1}^{n} x_i P(x_i)$$

$$\mu = \int_{-\infty}^{\infty} x f(x) dx$$

- **mean is defined similarly to discrete variables**

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 P(x_i)$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

- **variance is defined similarly to discrete variables**

Probability Density Function - Example

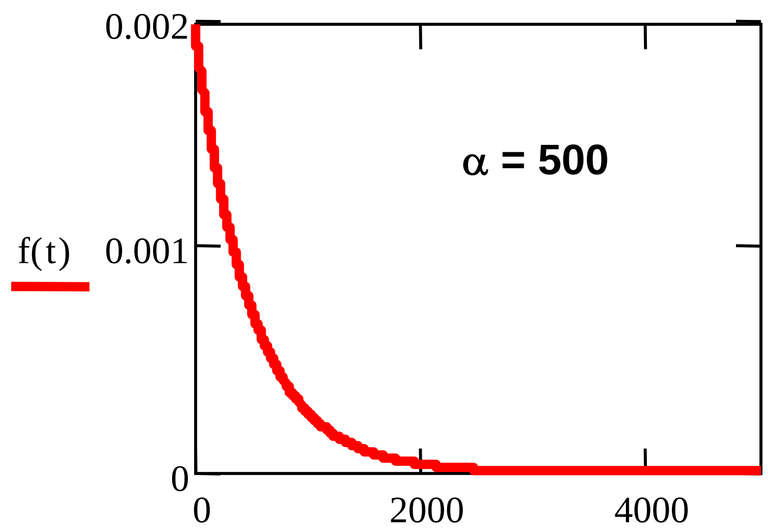
- Electrical component lifetime:

Probability that component with average lifetime α fails before time T:

$$P(t \le T) = \int_0^T f(t)dt$$

$$f(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{\alpha} e^{-\frac{t}{\alpha}} & t \ge 0 \end{cases}$$

$$P(t \le T) = \int_0^T \frac{1}{\alpha} e^{-\frac{t}{\alpha}} dt$$



Probability Density Function - Example

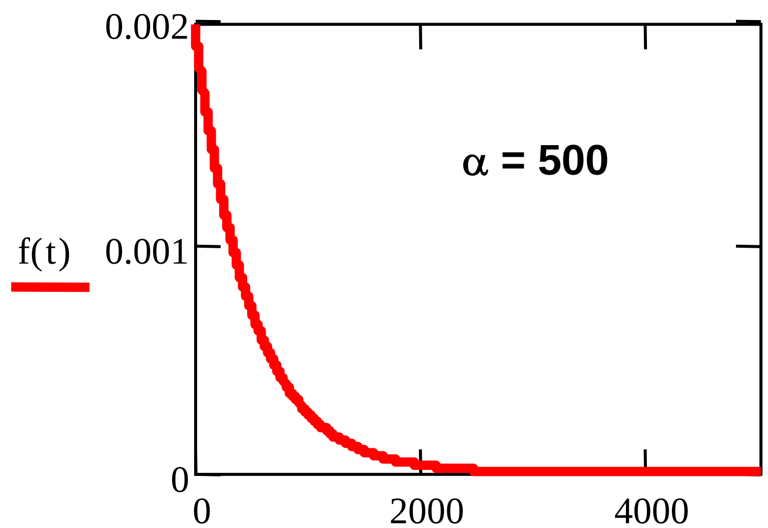
- Electrical component lifetime:

Probability that component with average lifetime α fails before time T:

$$P(t \le T) = \int_0^T f(t)dt$$

$$f(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{\alpha} e^{-\frac{t}{\alpha}} & t \ge 0 \end{cases}$$

$$P(t \le T) = \int_0^T \frac{1}{\alpha} e^{-\frac{t}{\alpha}} dt$$



What is the probability that a light bulb with a 500 hour expected lifetime will fail within the first 100 hours?

Probability Density Function - Example

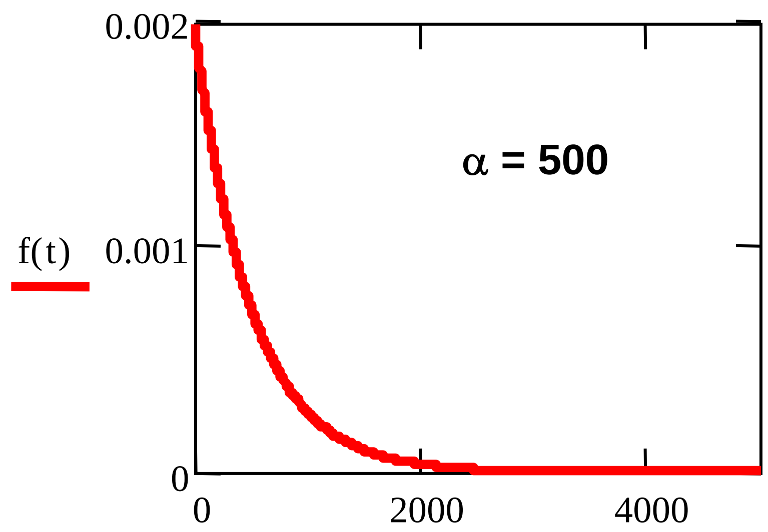
- Electrical component lifetime – exponential distribution:

Probability that component with average lifetime α fails before time T:

$$P(t \le T) = \int_0^T f(t)dt$$

$$f(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{\alpha} e^{-\frac{t}{\alpha}} & t \ge 0 \end{cases}$$

$$P(t \le T) = \int_0^T \frac{1}{\alpha} e^{-\frac{t}{\alpha}} dt$$



What is the probability that a light bulb with a 500 hour expected lifetime will fail within the first 100 hours?

$$\int_0^{100} \frac{1}{500} \cdot e^{\frac{t}{500}} dt = 0.221$$

Cumulative Distribution Function

Continuous random variable:

$$P(rv \le x) = \int_{-\infty}^{x} f(t)dt \doteq F(x)$$

Discrete random variable:

$$P(rv \le x_i) = \sum_{j=1}^{i} P(x_j)$$

Cumulative Distribution Function

Continuous random variable:

$$P(rv \le x) = \int_{-\infty}^{x} f(t)dt \doteq F(x)$$

Discrete random variable:

$$P(rv \le x_i) = \sum_{j=1}^{i} P(x_j)$$

Properties of C.D.F

$$P(a < x \le b) = F(b) - F(a)$$

$$P(x > a) = 1 - F(a)$$

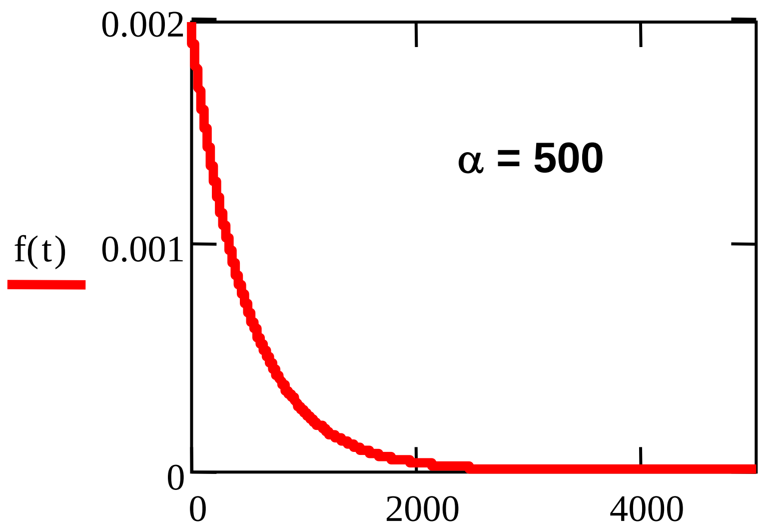
Exponential Distribution

- Probability Density Function

$$P(t \le T) = \int_0^T f(t)dt$$

$$f(t) = \begin{cases} 0 & t < 0 \\ \dfrac{1}{\alpha} e^{-\frac{t}{\alpha}} & t \ge 0 \end{cases}$$

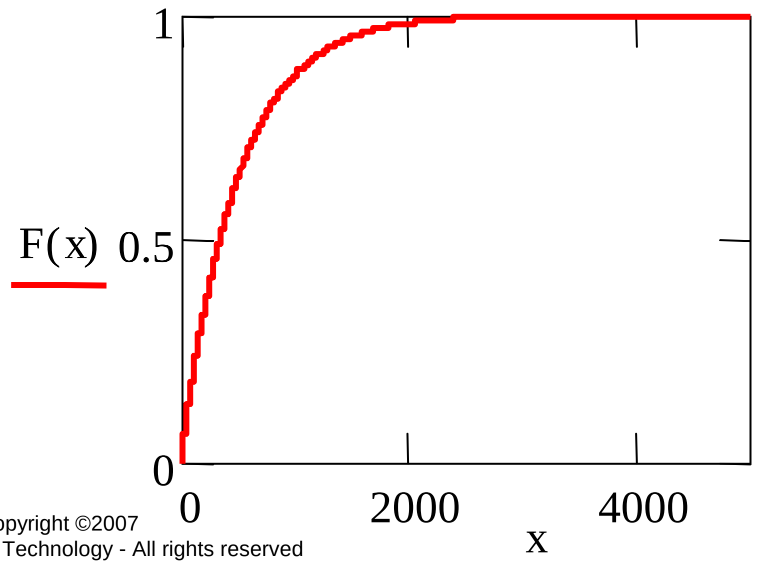
$$P(t \le T) = \int_0^T \frac{1}{\alpha} e^{-\frac{t}{\alpha}} dt$$



- Cumulative Distribution Function

$$P(rv \le x) = \int_{-\infty}^{x} f(t)dt \doteq F(x)$$

$$F(x) = \int_0^x \frac{1}{\alpha} e^{-\frac{t}{\alpha}} dt = 1 - e^{-\frac{x}{\alpha}}$$



Useful P.D.F.'s – Binomial Distribution

- A set of discrete random variables can have two possible outcomes: "success" or "failure"
  - A number of trials are performed, each "succeeds" or "fails"
  - P(success) remains constant during trials
  - n independent trials

- What is the probability of having exactly r successes?

Useful P.D.F.'s – Binomial Distribution

- A set of discrete random variables can have two possible outcomes: "success" or "failure"
  - A number of trials are performed, each "succeeds" or "fails"
  - P(success) remains constant during trials
  - n independent trials

- What is the probability of having exactly r successes?

$$P(r) = \binom{n}{r} p^r (1-p)^{n-r}$$

$$\binom{n}{r} \doteq \frac{n!}{r!(n-r)!}$$

Useful P.D.F.'s – Binomial Distribution

- A machining process creates a part with a dimension that has a random component. 85% of the parts lie within the required range of dimensions, but 15% are out of spec. What is the probability that 80 of 100 randomly chosen components will be acceptable?

Useful P.D.F.'s – Binomial Distribution

- A machining process creates a part with a dimension that has a random component. 85% of the parts lie within the required range of dimensions, but 15% are out of spec. What is the probability that 80 of 100 randomly chosen components will be acceptable?

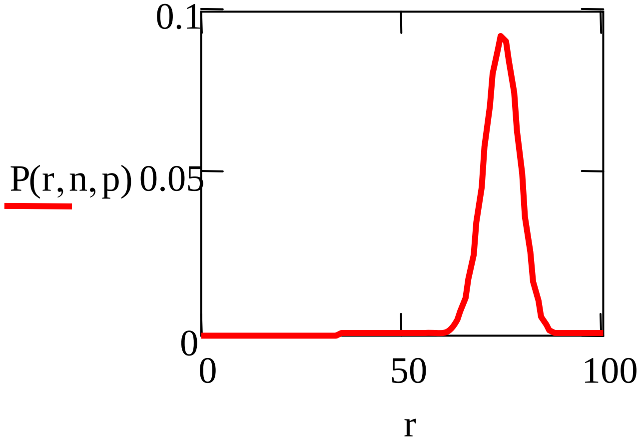
$$P(r) = \binom{n}{r} p^r (1-p)^{n-r} = \left(\frac{n!}{r!(n-r)!} \right) p^r (1-p)^{n-r}$$

$$P(r) = \left(\frac{100!}{80!(100-80)!} \right) (0.85)^{80} (0.15)^{100-80} = .04$$

Binomial Distribution

- Probability Distribution Function

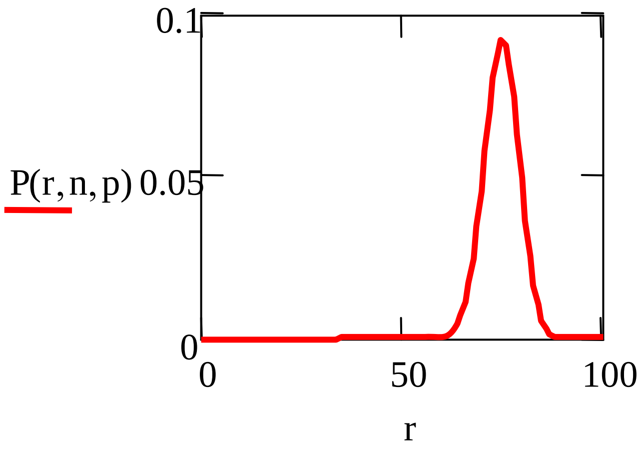
$$P(r) = \left(\frac{n!}{r!(n-r)!} \right) p^r (1-p)^{n-r}$$



Binomial Distribution

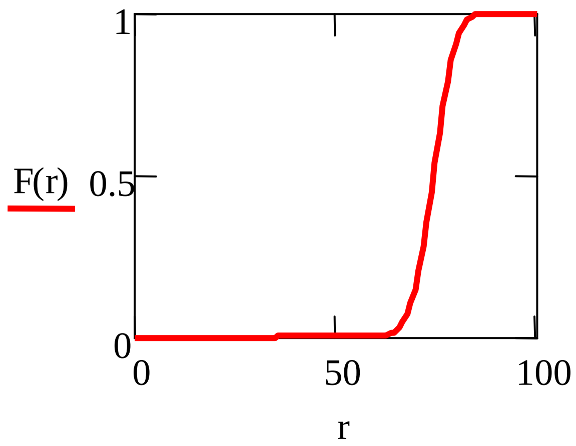
- Probability Distribution Function

$$P(r) = \left(\frac{n!}{r!(n-r)!} \right) p^r (1-p)^{n-r}$$



- Cumulative Distribution Function

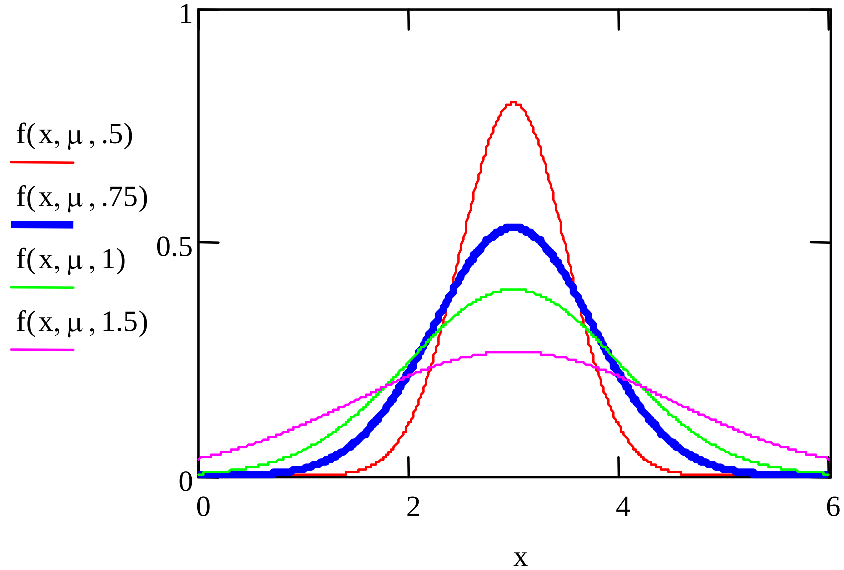
$$F(r) = \sum_{i=0}^{r} \left(\frac{n!}{i!(n-i)!} \right) p^i (1-p)^{n-i}$$



Gaussian Distribution

- Probability density function

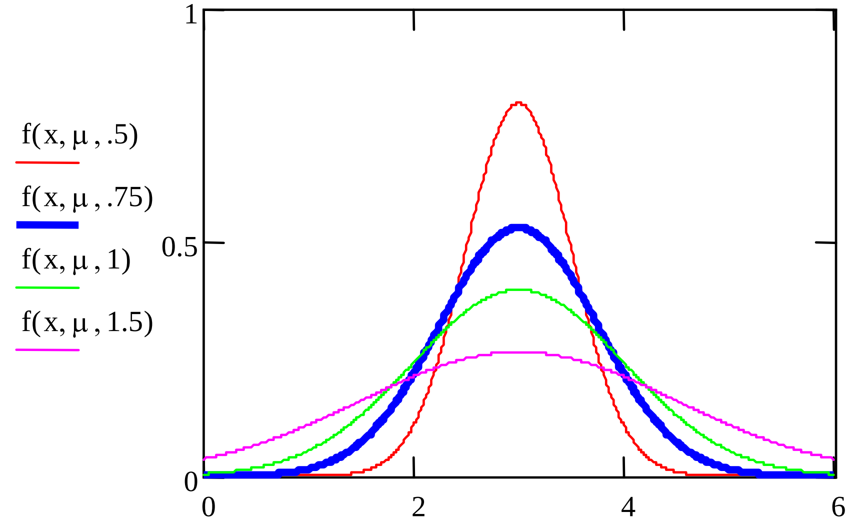
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



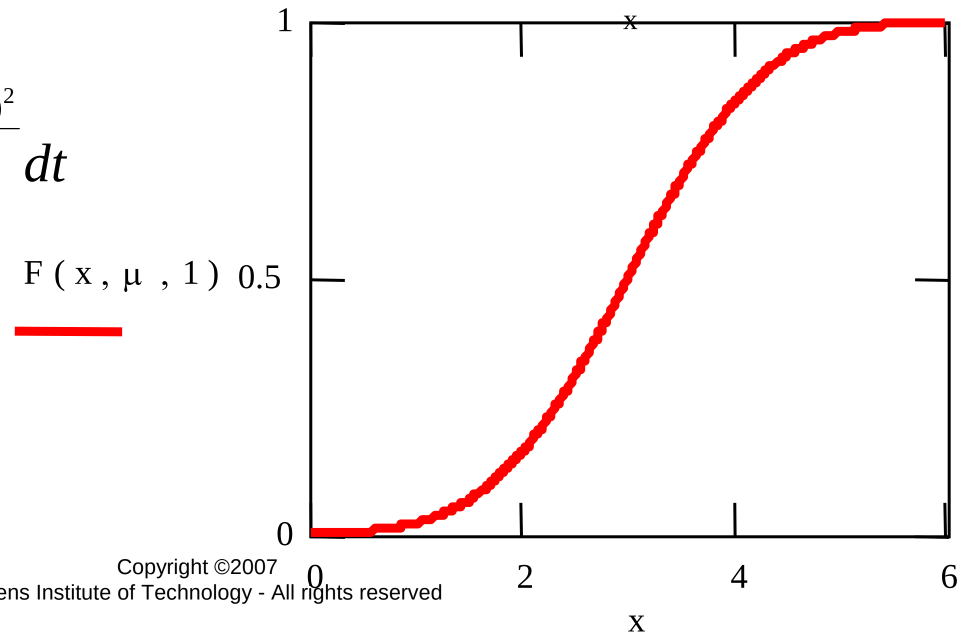
Gaussian Distribution

- Probability density function

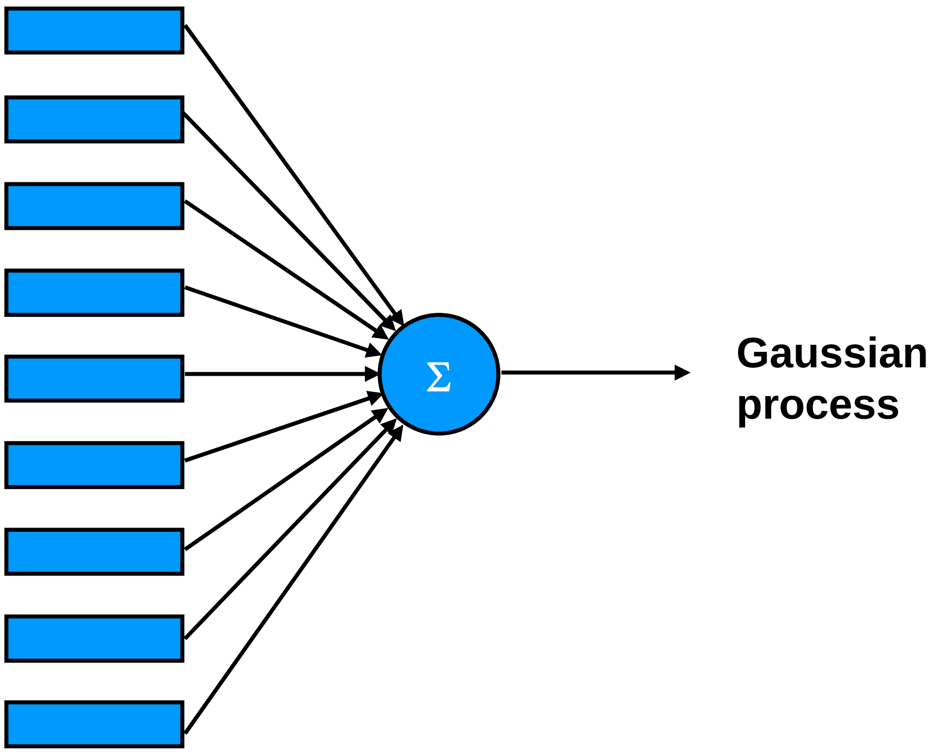
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



$$F(x) = \int_{-\infty}^{t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$



Source of Gaussian Distribution



Large number of independent, identically distributed r.v.s

Poisson Distribution

- A series of arrival events occur independently
- The average number of arrivals in a given duration interval is a constant

Poisson Distribution

- A series of arrival events occur independently
- The average number of arrivals in a given duration interval is a constant

- For an average arrival rate λ, the probability of x arrivals

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Poisson Distribution

- A series of arrival events occur independently
- The average number of arrivals in a given duration interval is a constant

- For an average arrival rate λ, the probability of x arrivals

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

- Probability that up to k arrivals will occur:

$$P(x \le k) = \sum_{i=0}^{k} \frac{e^{-\lambda}\lambda^{i}}{i!}$$

Poisson Distribution

- Software code reliability is sometimes measured in FITs – Failures In a Thousand lines of code
- Assume that Microsoft Vista (10’s of millions of lines of code) has a FIT rate of .5
- What is the probability that a 10,000 line subroutine has 2 or more errors?

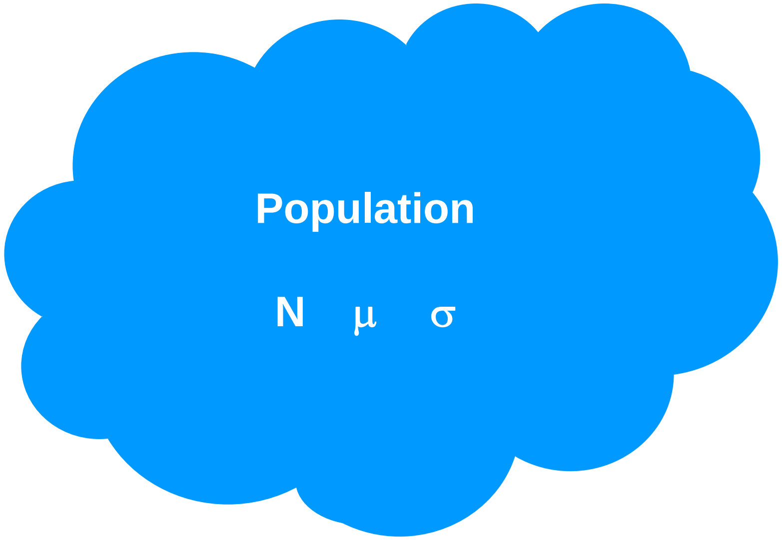
Poisson Distribution

- Software code reliability is sometimes measured in FITs – Failures In a Thousand lines of code
- Assume that Microsoft Vista (10's of millions of lines of code) has a FIT rate of .5
- What is the probability that a 10,000 line subroutine has 2 or more errors?

$$P(x \le k) = \sum_{i=0}^{k} \frac{e^{-\lambda} \lambda^{i}}{i!}$$

$$P(x \ge 2) = 1 - P(x \le 1) = 1 - \sum_{i=0}^{1} \frac{e^{-(.5 \cdot 10)} (.5 \cdot 10)^{i}}{i!} = .09$$

Parameter Estimation

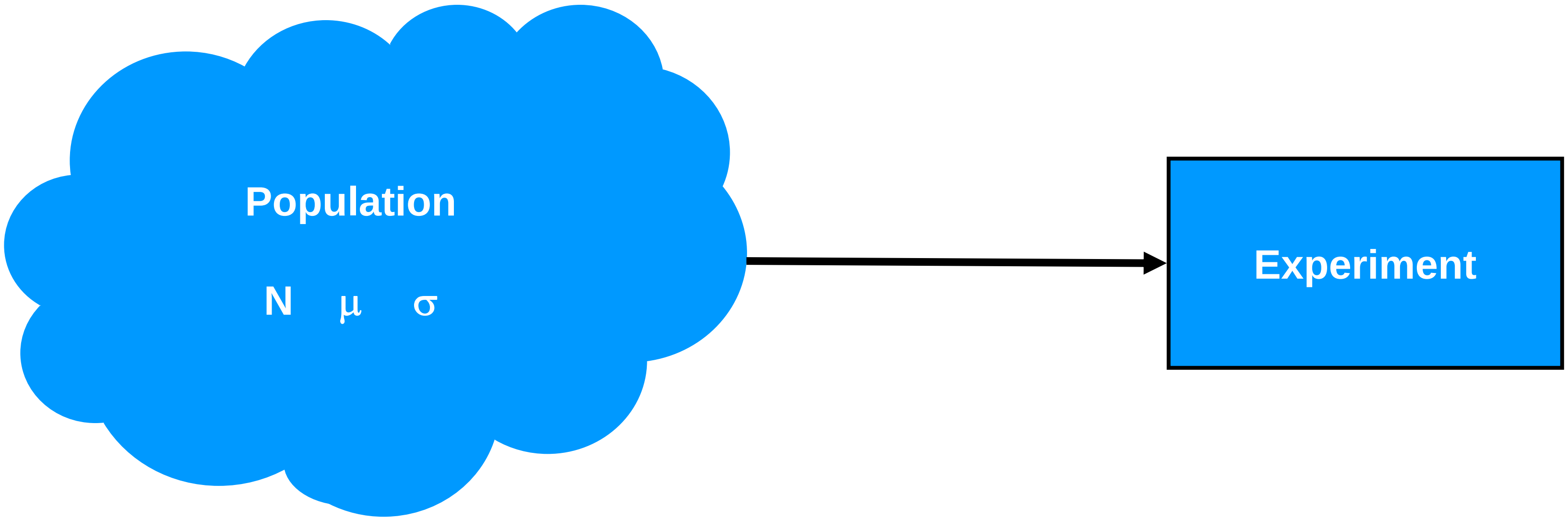


N potential outcomes

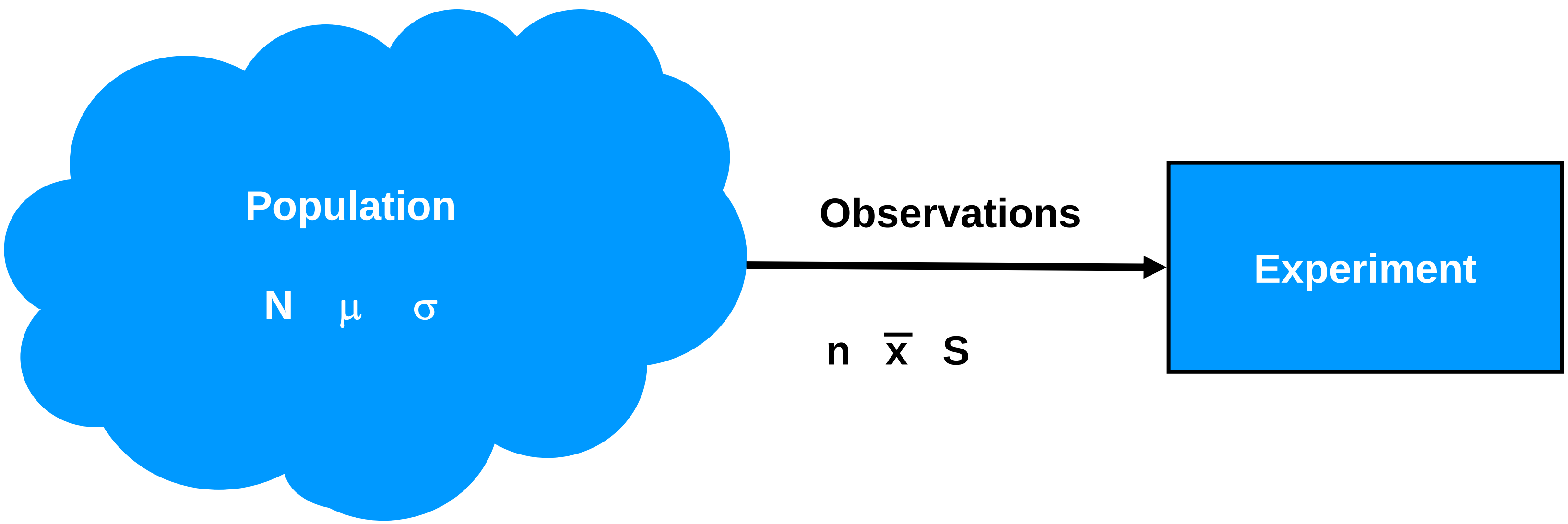
Mean value μ

Standard deviation σ

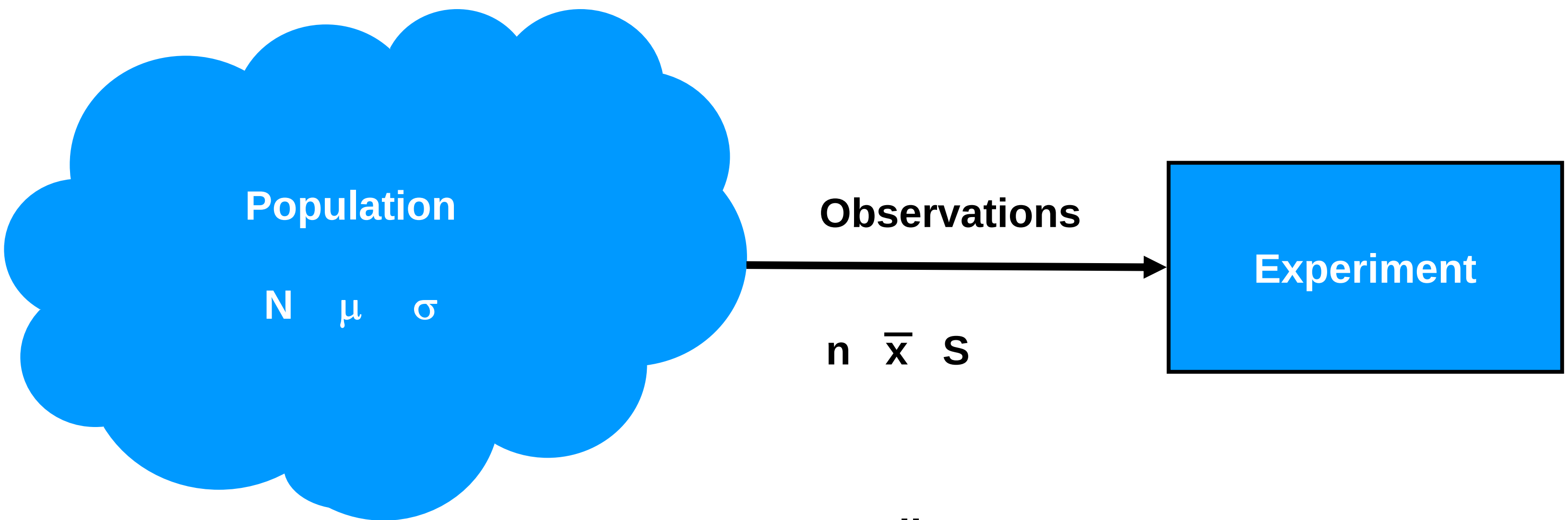
Parameter Estimation



Parameter Estimation

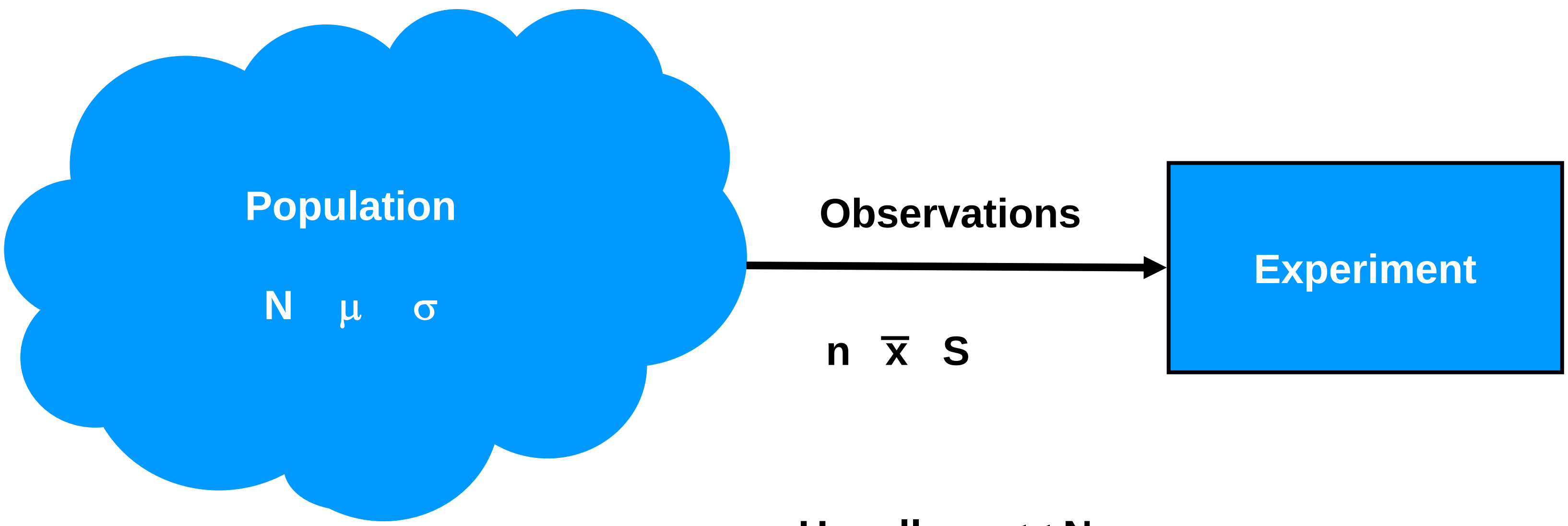


Parameter Estimation



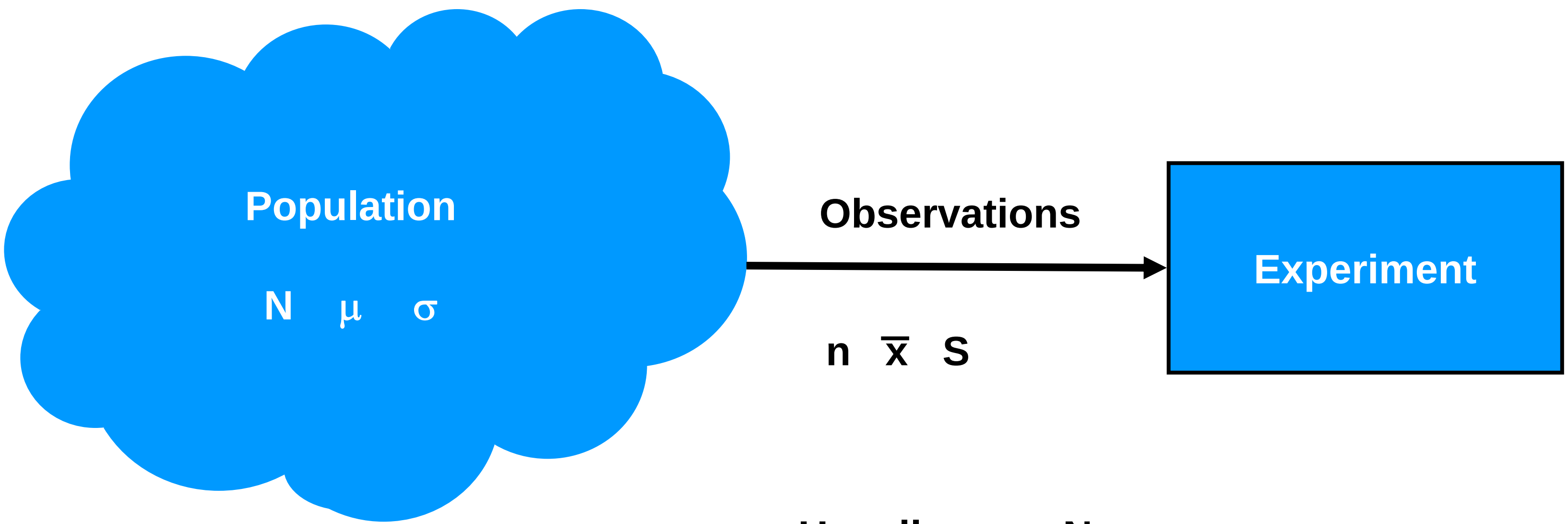
Usually, n << N

Parameter Estimation



Usually, n << N
so $\bar{x} \sim \mu$
$S \sim \sigma$

Parameter Estimation



Usually, $n << N$
so $\bar{x} \sim \mu$
$S \sim \sigma$

But if $\delta = |\bar{x} - \mu|$,
how large is δ likely to be?

Next time

- More On Statistical Analysis of Experimental Data
  - Confidence Intervals
  - Correlation