Section 2: Probability Theory

- *Purpose*
  - *Students will review fundamentals of probability*
  - *Become familiar with several probability distributions that are commonly encountered in PRA*
- *Objectives*
  - *Students will be able to calculate simple probabilities involving*
    - *"AND", "OR", "NOT" operations*
    - *Conditional probabilities, independent events*
    - *Bayes' theorem*
    - *Poisson, binomial, and exponential distributions*
  - *Students will understand the terms mean, variance, percentile, and be able to relate these to particular distributions used in the course*

Probability Theory

- *Topics to be covered include*
  - *Basic (Frequentist) Framework*
  - *Rules for Manipulating Probabilities*
  - *Discrete Probability Distributions*
  - *Continuous Probability Distributions*
  - *Moments, Percentiles*
  - *Relations among Distributions*

Basic (Frequentist) Framework

- *A repeatable experiment can result in a number of outcomes. Experiment may be "trial", "test", "demand", etc.*
- *Sample space **S** is the set of all possible outcomes on any one experiment*
- ***Probability** of any outcome is long-term fraction of times when the outcome occurs*
  - *Probability is "reached" after infinite series of these outcomes*
  - *Identical experiment must be repeated over and over*
- *An **event** is a set of outcomes*
  - *Its probability is the sum of the probability of each constituent outcome*

HOPE

Pages A-1 through A-4

Example 1

- *Experiment: Try to start EDG-A*
- *The possible outcomes (i.e. the sample space, S)*
  - *Failure to start (FTS_A)*
  - *Start but failure to run (FTR_A)*
  - *Start and run to end of mission ($Success_A$)*
- *Some possible events*
  - *EDG-A fails somehow*
  - *EDG-A starts*
  - *Etc.*

Example 2

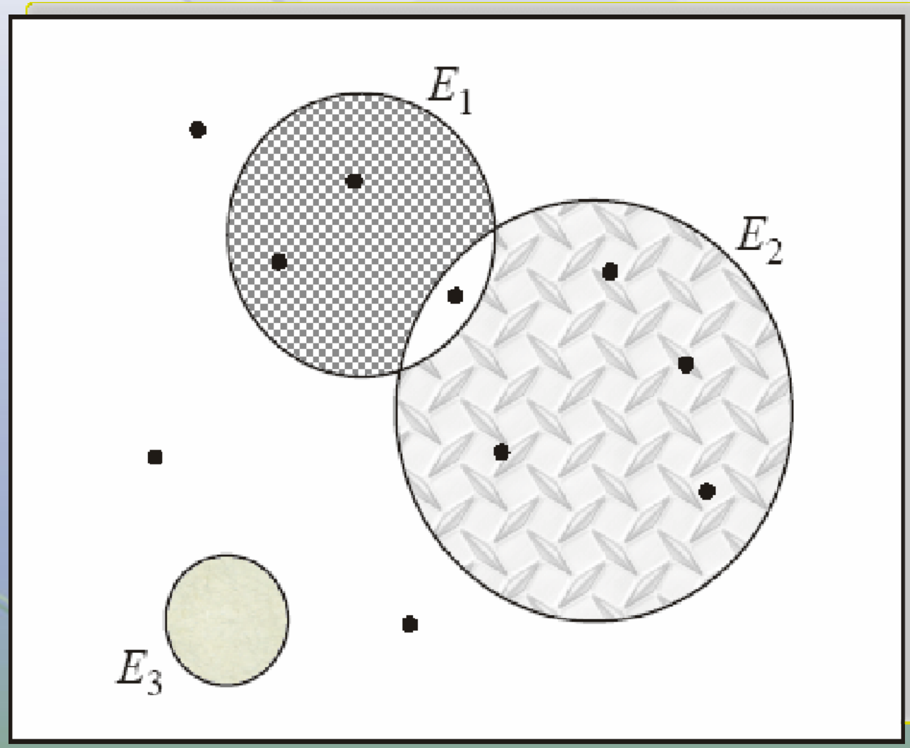
- Experiment: Try to start two EDGs, EDG-A and EDG-B
- The outcomes (i.e. the sample space):

FTS_A & FTS_B	FTS_A & FTR_B	FTS_A & $Success_B$
FTR_A & FTS_B	FTR_A & FTR_B	FTR_A & $Success_B$
$Success_A$ & FTS_B	$Success_A$ & FTR_B	$Success_A$ & $Success_B$

- Some possible events
  - At least one EDG succeeds
  - Both EDGs fail somehow
  - At least one EDG fails to start
  - Exactly one EDG fails
  - Etc.

Example 3

- *It is sometimes helpful to show* ***events*** *and* ***outcomes*** *via a Venn diagram*
  - *Three events, 10 outcomes*



Building Events from Other Events or Outcomes — OR

- *A OR B = event containing all outcomes that are in A or in B*
  - *Also written A∪B, the **union** of A and B*
  - *The union symbol, ∪, is easy to remember since symbol looks like the letter "U"*
- *In a PRA, minimal cut sets are "ORed" together to obtain overall results of the analysis*

Building Events from Other Events or Outcomes — AND

- *A AND B = event containing all outcomes that are both in A and in B*
  - *Also call the **intersection** of A and B, written A∩B*
  - *The intersection symbol ∩ can be remembered as the opposite of the union symbol*
- *In a PRA, the events within a **single** minimal cut set are "ANDed" together to obtain the cut set value*
- *A and B are **disjoint** or **mutually exclusive** if they have no outcomes in common, i.e. A AND B is empty*

Building Events from Other Events or Outcomes — NOT

- *The **complement** of A, or NOT A, is the event containing all the outcomes that are not in A.*
- *Sometimes written* $\overline{A}$ *or /A*

Basic (Frequentist) Framework (cont.)

- Within the Frequentist approach to statistics, there are two divisions of thought
  - Relative frequency
    - Repeated (unchanging) experiment yields measurable outcomes
    - Event probability achieved in limit of many trials
      - Example, toss die 100,000 times
      - 17,003 times “1” seen…Pr(1) ≅ 0.170
  - Classical (Bernoulli) approach
    - Equal probable outcomes
    - Event probability available directly by ratio of outcomes of interest versus all possible outcomes
      - Example, Pr(1 | roll of die) = 1 / 6 = 0.1666…

Tips for Solving Probability Problems

- *Write what you have*
  - *Can you list the outcomes?*
  - *What events are relevant?*
  - *What is “fixed” and what is “random”?*
  - *What is the problem asking for?*
  - *What formulas relate to this question?*
- *Do not try to do everything in your head.*
  - *Use pencil and paper, and proceed step by step through the problem*

Elementary "Rules" of Probability

- Each event has an associated probability
- Probability, or p, is a real value
  - Possible values are bounded by (and include) 0.0 and 1.0
  - Probability of union of non-overlapping (disjoint) events is the sum of the event probabilities
  - Probability of all possible outcomes (i.e., the sample space) equal to 1.0

Rules for Manipulating Probabilities

- *The **NOT** (or complement) operation*
  - *Subtract probability from 1.0*
  - *Example, Pr(not A) = 1 - Pr(A)*

- *Another probability problem tip*
  - *With messy problems using terms such as “at least” or “at most,” try to calculate complement of the probability first*
    - *Pr(A) = 1 - Pr(not A)*

Rules for Manipulating Probabilities (cont.)

- *The **OR** (or union) operation*
  - *If A, B are **disjoint***
    - *Pr(A or B) = Pr(A) + Pr(B)*
  - *If A, B are not disjoint*
    - *Pr(A or B) = Pr(A) + Pr(B) - Pr(A AND B)*

Rules for Manipulating Probabilities (cont.)

- *The **AND** (or intersection) operation*
  - *If A, B are **independent***
    - *Pr(A AND B) = Pr(A) • Pr(B)*
  - *If A, B are not independent (i.e., dependent, or **conditional**)*
    - *Pr(A AND B) = Pr(A) • Pr(B | A)*

      = Pr(B) • Pr(A | B)

Rules for Manipulating Probabilities (cont.)

- ***Conditional*** *probability definition*
  - *We said that in general*
    - *Pr(A AND B) = Pr(A) • Pr(B | A)*
  - *The conditional probability is last term Pr(B | A), so*
    - *Pr(B | A) = Pr(A AND B) / P(A)*
    - *Pr(A | B) = Pr(A AND B) / P(B)*
  - *These last equations define "conditional probability".*
- *We will see that this product rule of conditional probabilities leads us to "Bayes' Theorem"*

Do Not Confuse Independent and Disjoint

- *If A, B are mutually exclusive (i.e., disjoint), then*
  - *Pr(A AND B) = 0*
- *If Pr(A AND B) = 0 but Pr(A) • Pr(B) ≠ 0*
  - *Thus, A and B are **not** independent*
  - *Instead, they simply are disjoint*
    - *On a Venn diagram, they do not overlap*

Bayes' Theorem

- A set of events $\{C_i\}$ is a **partition** of the sample space
  - If C_is are mutually exclusive
    - Each pair is mutually exclusive…no overlap
  - And if union of C_is is the entire sample space
- Bayes' Theorem: If $\{C_i\}$ is a partition of the sample space,

$$\Pr(C_i \mid E) = \frac{\Pr(E \mid C_i)\Pr(C_i)}{\sum \Pr(E \mid C_j)\Pr(C_j)}$$

- Bottom term is Pr(E) (where E is the "evidence")

$$\Pr(E) = \sum \Pr(E \mid C_j)\Pr(C_j)$$

is called "Law of Total Probability"

HOPE

Pages A-4 through A-12

Bayes' Theorem

- *If we are calculating probability of event C where evidence E is available*
  - *Pr(C | E) = Pr(C) Pr(E | C) / Pr(E)*
- *Terms in equation above have specified names*

Pr(C | E)	*Posterior probability*
Pr(C)	*Prior probability*
Pr(E | C)	*Likelihood*
Pr(E)	*Unconditional probability of evidence*

Bayes Example

- *Tests for integrity are carried out on manufactured radiation sources*
  - *Manufacturer A results: 10% cracked, over long run*
  - *Manufacturer B results: 5% cracked, over long run*
- *Hospital gets 60% of its sources from manufacturer B*
- *Incident report is later sent to the NRC stating leak from a source at a hospital due to cracked source*
  - *What is the probability that the hospital used manufacturer B?*
  - $P(\text{Manufacturer B}|\text{crack}) = (0.6)(0.05)/[(0.4)(0.10) + (0.6)(0.05)]$

    $= 0.20$

Second Bayes' Theorem Example

- *Look at an example…the infamous "game show" problem*
- ***One** of three doors contains a prize. You select door #1.*
  - *The host opens door #2 (which contains nothing) and asks you if would like to **switch** to door #3*
  - *What does Bayes' Theorem say?*
    - *"C_1" in this case is the event our door, #1, hides the prize*
    - *The evidence, E, is knowledge that door #2 does not have prize*
- *Let us first see what information we have*
  - *$Pr(C_1) = Pr(C_2) = Pr(C_3)$ = prior probability door holds the prize independent of any new information = 1/3*
  - *$Pr(E \mid C_1)$ = conditional probability that door #2 would have been opened by host = ½*
  - *$Pr(E \mid C_2) = 0$, and $Pr(E \mid C_3) = 1$ — The host knows where the prize is, and will never open that door!*

Second Bayes' Theorem Example (cont.)

- $Pr(C_1 \mid E)$ = probability that our door has the prize, given door #2 was opened

  $= Pr(C_1)\ Pr(E \mid C_1) / Pr(E)$ **Bayes!**

- We know $Pr(C_1)$ and $Pr(E \mid C_1)$, but what is $Pr(E)$?
  - Use Law of Total Probability
- Pr(E) = Pr(#1 has prize) Pr(#2 opened | #1 has prize) +
  Pr(#2 has prize) Pr(#2 opened | #2 has prize) +
  Pr(#3 has prize) Pr(#2 opened | #3 has prize)
  = (1/3)(1/2) + (1/3)(0) + (1/3)(1) = 1/2
- We now have all the parts, doing the math indicates
  - $Pr(C_1 \mid E) = (1/3)(1/2)/(1/2) = 1/3$
  - $Pr(C_3 \mid E) = (1/3)(1)/(1/2) = 2/3$

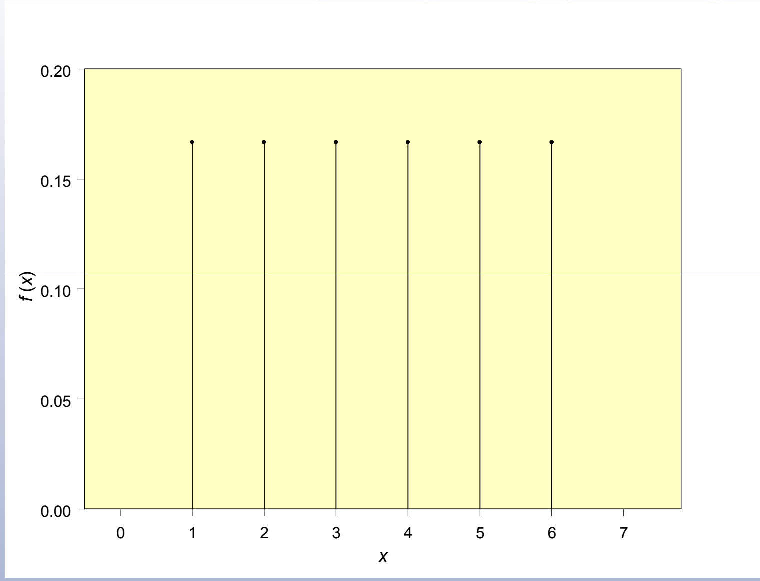
Discrete Probability Distributions

- *Outcomes can be summarized by a **random** variable X, which takes possible values x*
- *Probabilities are based on X's probability distribution function (pdf)*
  - $f(x) = Pr(X = x)$
- *Facts about a pdf*
  - $f(x_i) \geq 0$
  - $SUM[\, f(x_i)\,]_{all\ i} = \sum f(x_i) = 1$

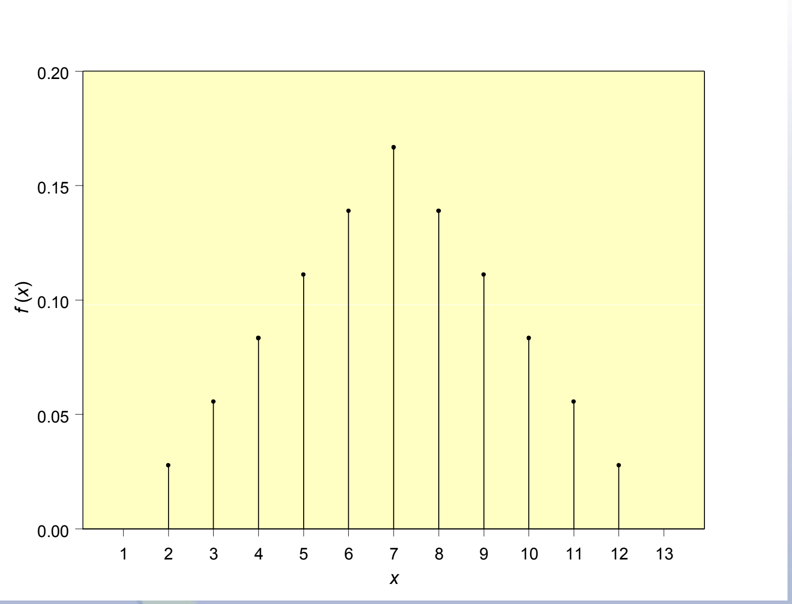
HOPE

Page A-5

Examples: Number of Spots on Dice



Spots on One Die



Total Spots on Two Dice

Random?

- Wolfram's thought of randomness in "New Kind of Science"
  - Try standard simple programs to detect regularities/patterns
  - If none detected, likely no other tests will show nonrandom behavior
  - He does not consider something to be truly random if generated from simple rules
- In other words
  - "… one considers a sequence 'random' if no patterns can be recognized in it, no predictions can be made about it, and no simple description of it can be found."

Random?

- *Consulting a dictionary…*

  "lacking any definite plan or order or purpose; governed by or depending on chance"
- *Jaynes' thought is that there is no such thing as "random," including no such thing as random variables*
  - *We only have variables whose values may not be known with certainty*

Two Most Popular Discrete Distributions

- *Poisson*
- *Binomial*

Poisson Distribution

- *Most commonly used model for initiating events.*
- *Assumptions on the physical process*
  1. *Probability of event in short time period Δt is approximately $\lambda \times \Delta t$, for some constant λ*
  2. *Simultaneous events do not occur*
  3. *Occurrences of events in disjoint time periods are independent*
- *Assumption on the data*
  4. *We observe a random number of events X in a fixed time period t*

HOPE

Pages 2-2 through 2-7, A-13 through A-14

Poisson Distribution (cont.)

- *Then X has a **Poisson**(λt) distribution:*

$$\Pr(X = x) = e^{-\lambda t}(\lambda t)^x / x!$$ *, for x = 0, 1, 2, …*

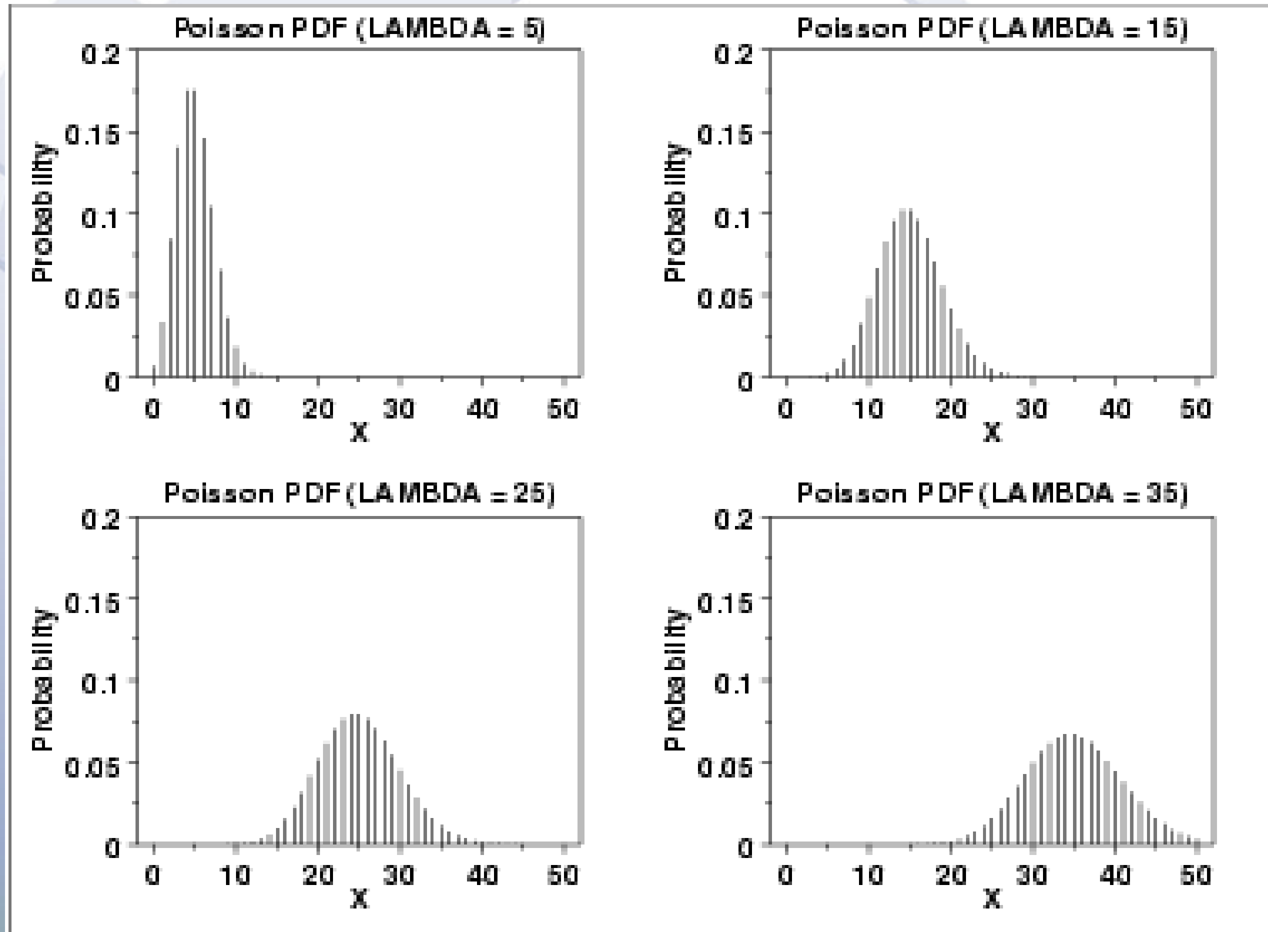
(x = number failures)

(for pictures, HOPE page A-14)

- *The distribution depends on one quantity, λt*
  - *Therefore, product λt is sometimes written as μ, and the distribution is called Poisson(μ)*
- *Moments*
  - *Mean = λt*
  - *Variance = λt*

Poisson Distribution (cont.)

- *L' image*



http://www.itl.nist.gov/div898/handbook/eda/section3/eda366.htm

Binomial Distribution

- *One commonly used model for failure to start.*
- *Assumptions on the physical process*
  1. *On each demand, outcome is a failure with probability p (success with probability 1 – p)*
     - *This p is the same for all demands*
  2. *Occurrences of failures on different demands are independent*
- *Assumption on the data*
  3. *We observe a random number of failures X in a fixed number of demands n*

HOPE

Pages 2-7 through 2-10, A-12 through A-13

Binomial Distribution (cont.)

- *Then X has a* **binomial***(n, p) distribution:*

$$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$ *, for x = 0, 1, …, n*

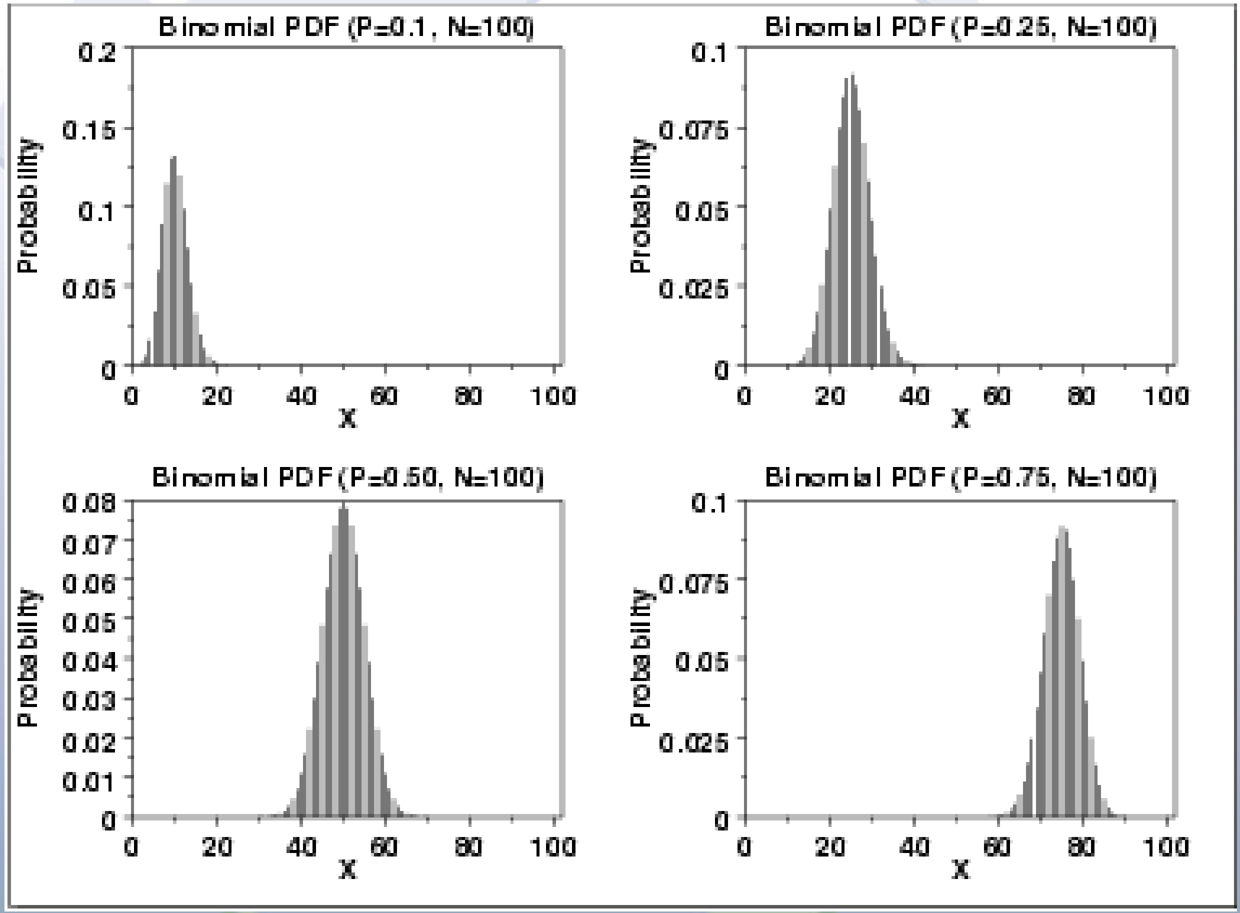
(x = number failures)

(for pictures, HOPE page A-13)

- *Moments*
  - *Mean = n p*
  - *Variance = n p (1 – p)*

Binomial Distribution (cont.)

- *L' image*



http://www.itl.nist.gov/div898/handbook/eda/section3/eda366.htm

Continuous Probability Distributions

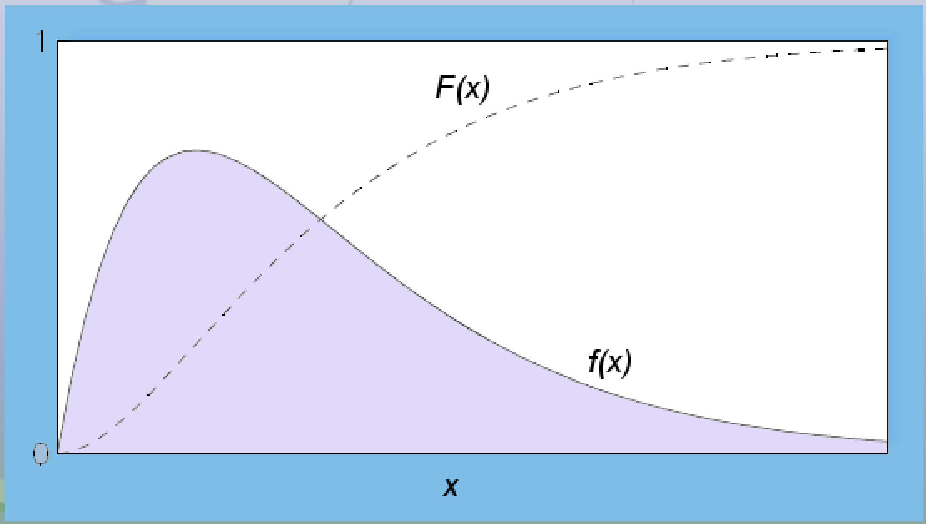
- *Random variable X takes values in a continuous range, such as from 0 to ∞*
- *In principle, Pr(a ≤ X ≤ b) =* $\int_a^b f(x)dx$,
  - *where f(x) is the probability density function.*
- *In practice, Pr(a ≤ X ≤ b) = F(b) – F(a)*
  - *where F is the cumulative distribution function (cdf)*

HOPE

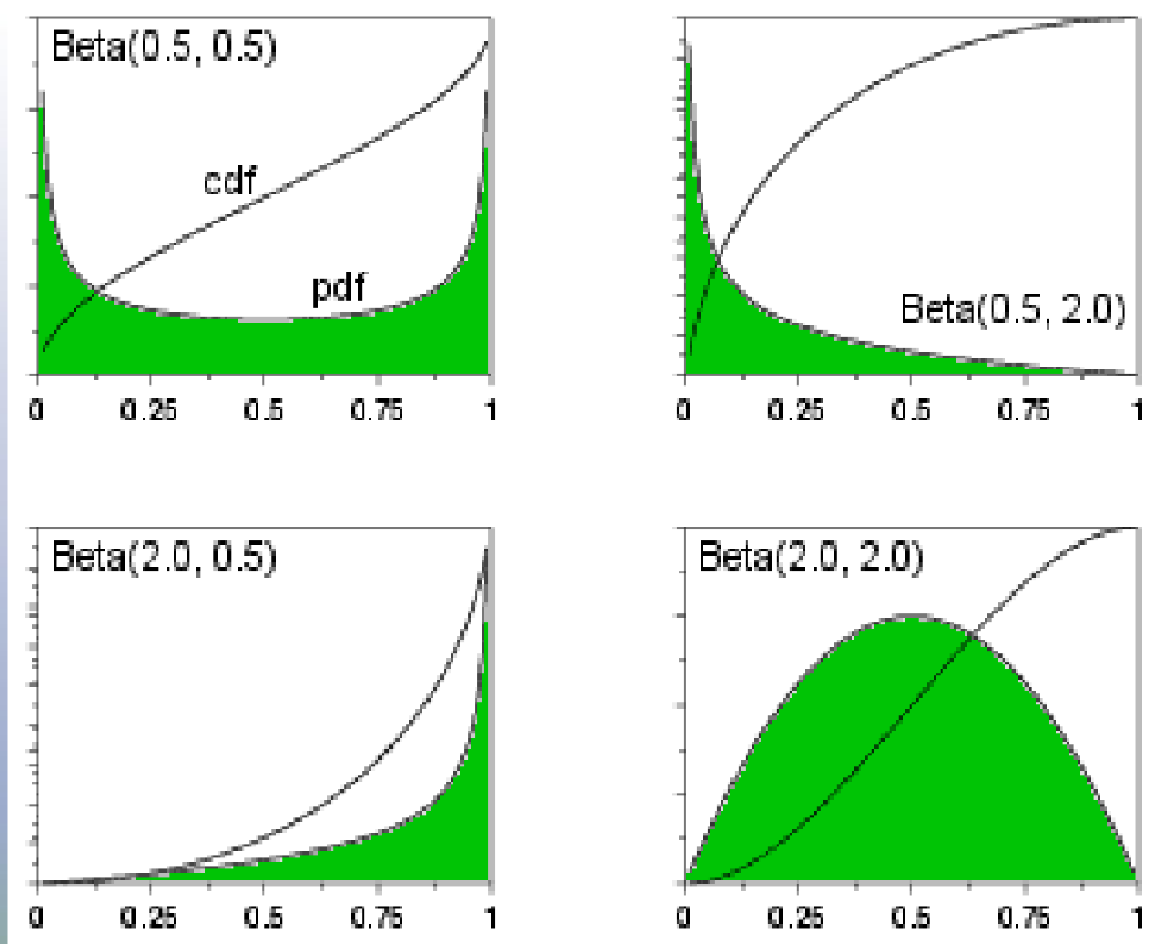
Pages A-5 through A-6

Continuous Probability Distributions (cont)

- Relations
  - $F(x) \equiv Pr(X \leq x) = \int_{-\infty}^{x} f(x')dx'$
  - $f(x) = dF(x)/dx$
- Note, Pr(X = x) = 0 for any number x
  - But probability that X is in an interval is typically nonzero



Continuous Probability Distributions (cont)



Exponential Distribution

- A commonly used distribution for a duration
- Very simple (sometimes too simple)
- Setting: Watch something until an event of interest occurs
  - Failure to run
  - Restoration of power
  - Suppression of fire, etc.
- Let T be time when event occurs

HOPE

Pages 2-17, A-17 through A-18

Exponential Distribution (cont.)

- *Assumption on the physical process*

  1. *For $t \geq 0$ and small Δt*

     $$Pr(T \leq t + \Delta t \mid T > t) \approx \lambda \times \Delta t \quad \text{(for a constant } \lambda)$$

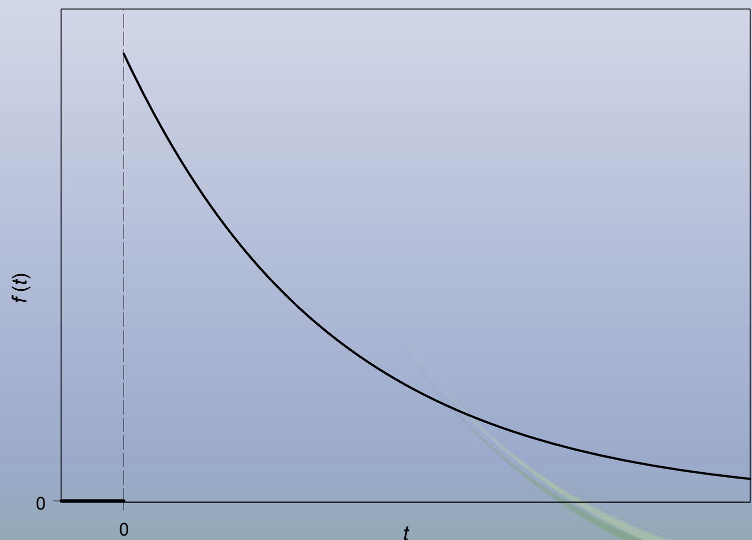
- *Interpretation*
  - *If the system is running at time t, probability that system will fail in next small time interval Δt is $\lambda \times \Delta t$, regardless of what t is.*
  - *That is, the system does not improve or age as a function of time*
- *Assumption on the data*

  2. *We observe the event time, T*

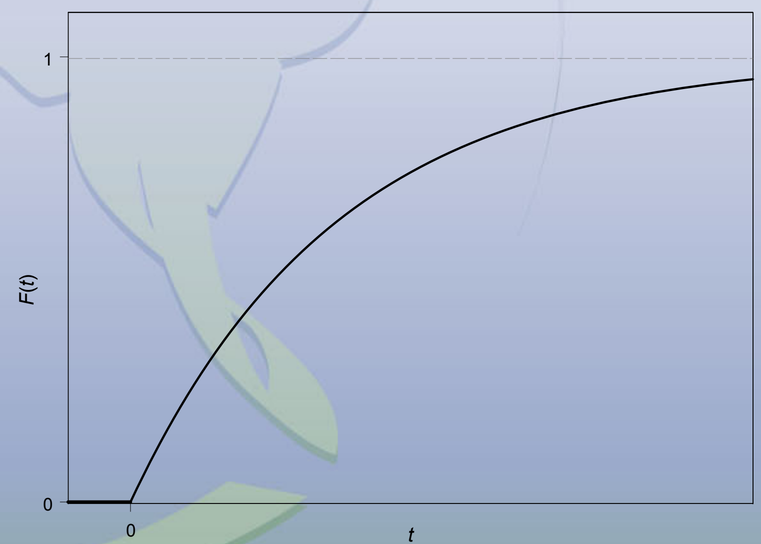
Exponential Distribution (cont.)

- *Under the above assumption*
  - *T has an* **exponential**(λ) *distribution*

$$f(t) = \lambda e^{-\lambda t} \quad \text{for } t \geq 0$$



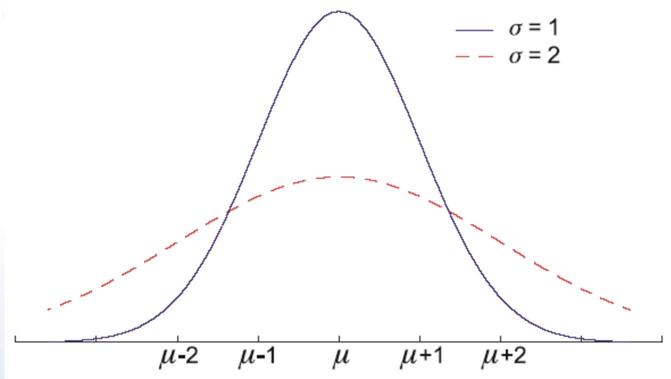
$$F(t) = 1 - e^{-\lambda t} \quad \text{for } t \geq 0$$



Exponential Distribution (cont.)

- Units
  - "λt" is unitless
  - λ has units of 1/t (in PRA, usually per hour)
    - Some initiating events are per year
- Alternative parameterization in terms of μ = 1/λ.
  - Just rewrite formulas in obvious way
  - Units of μ are units of t
    - Also known as "mean time to failure"

Normal Distribution



- *Arises in many settings*
  - *Primary application in this course is as a means to the lognormal distribution*
  - *Density equation in HOPE, page A-15*
- *If X has a* **normal**(μ, σ^2) *distribution, then*

$$\Pr(X \le x) = \Pr\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Φ *is tabulated in many books, for example HOPE Table C-1*

HOPE **Pages A-15 through A-16**

Moments and Percentiles

- *The **mean**, or **expected value**, or **expectation**, of X is weighted average of the values of X*
  - $E(X) = \sum_{x} x \Pr(X = x) = \sum_{x} x f(x)$ *if X discrete*
  - $E(X) = \int_{-\infty}^{\infty} x f(x) dx$ *if X continuous*

HOPE

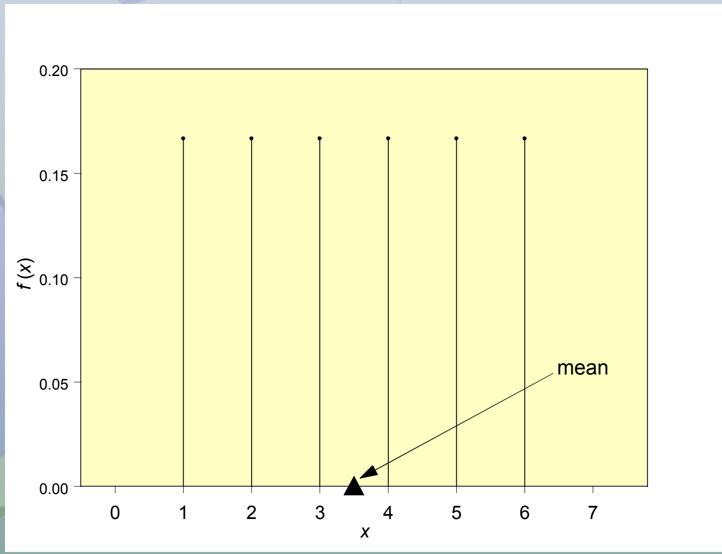
Pages A-8 through A-10

Moments and Percentiles (cont.)

- *The **variance** is the analogous weighted average of $[X - E(X)]^2$:*
  - $\mathrm{var}(X) = \sum_x [x - E(X)]^2 f(x)$ *if X discrete*
  - $\mathrm{var}(X) = \int_{-\infty}^{\infty} [x - E(X)]^2 f(x)dx$ *if X continuous*
- *The **standard deviation** is the square root of the variance*

Mean Example

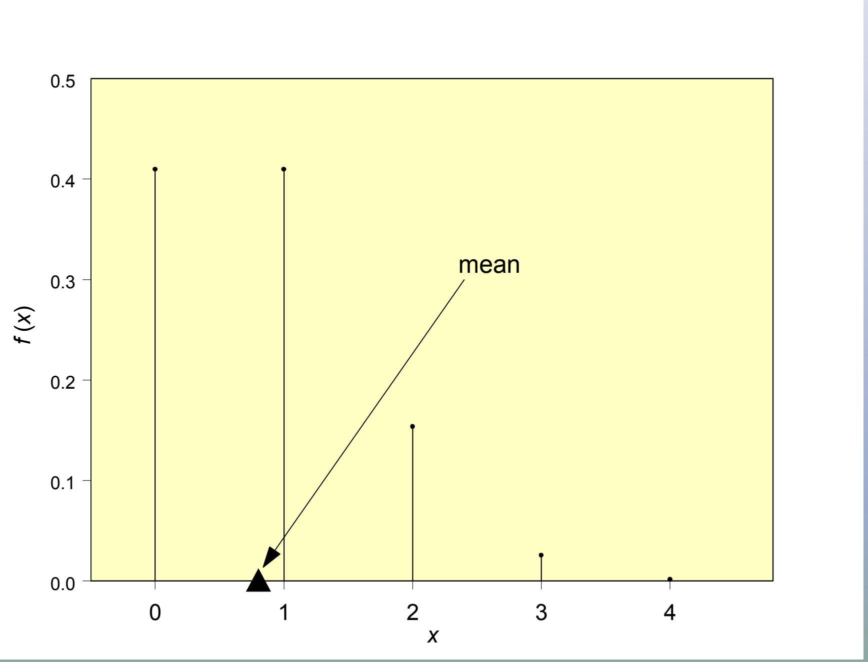
- We have the discrete distribution for a single die
  - What is the expected value?
  - $Pr(X = x) = 1/6$
  - $E[X] = 1(1/6) + 2(1/6) + 3(1/6) + 4(1/6) + 5(1/6) + 6(1/6)$
    $= 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 1 = 3.5$
  - Since discrete, we can not really get an outcome of 3.5.
    It has to be 1, 2, 3, 4, 5, or 6



Second Mean Example — Binomial(n=4, p=0.2)

- Possible values are 0, 1, 2, 3, 4.
- The distribution is skewed, not symmetrical
- Smaller values have higher probability, so mean is weighted toward small side
- Direct calculation shows

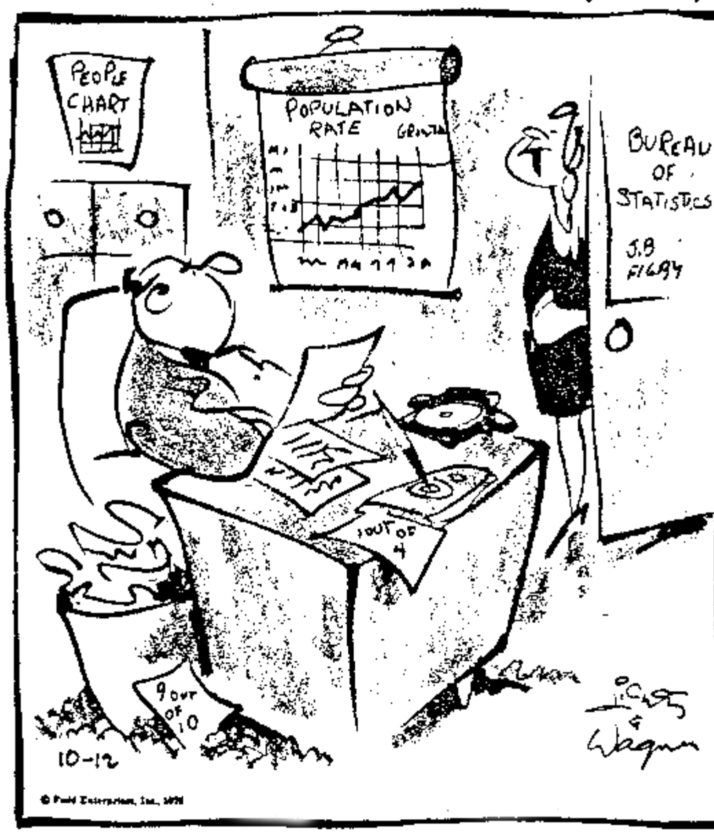
  mean = 0.8



Moments and Percentiles (cont.)

- *The distinction between the mean and the full distribution is the basis of most statistician jokes.*

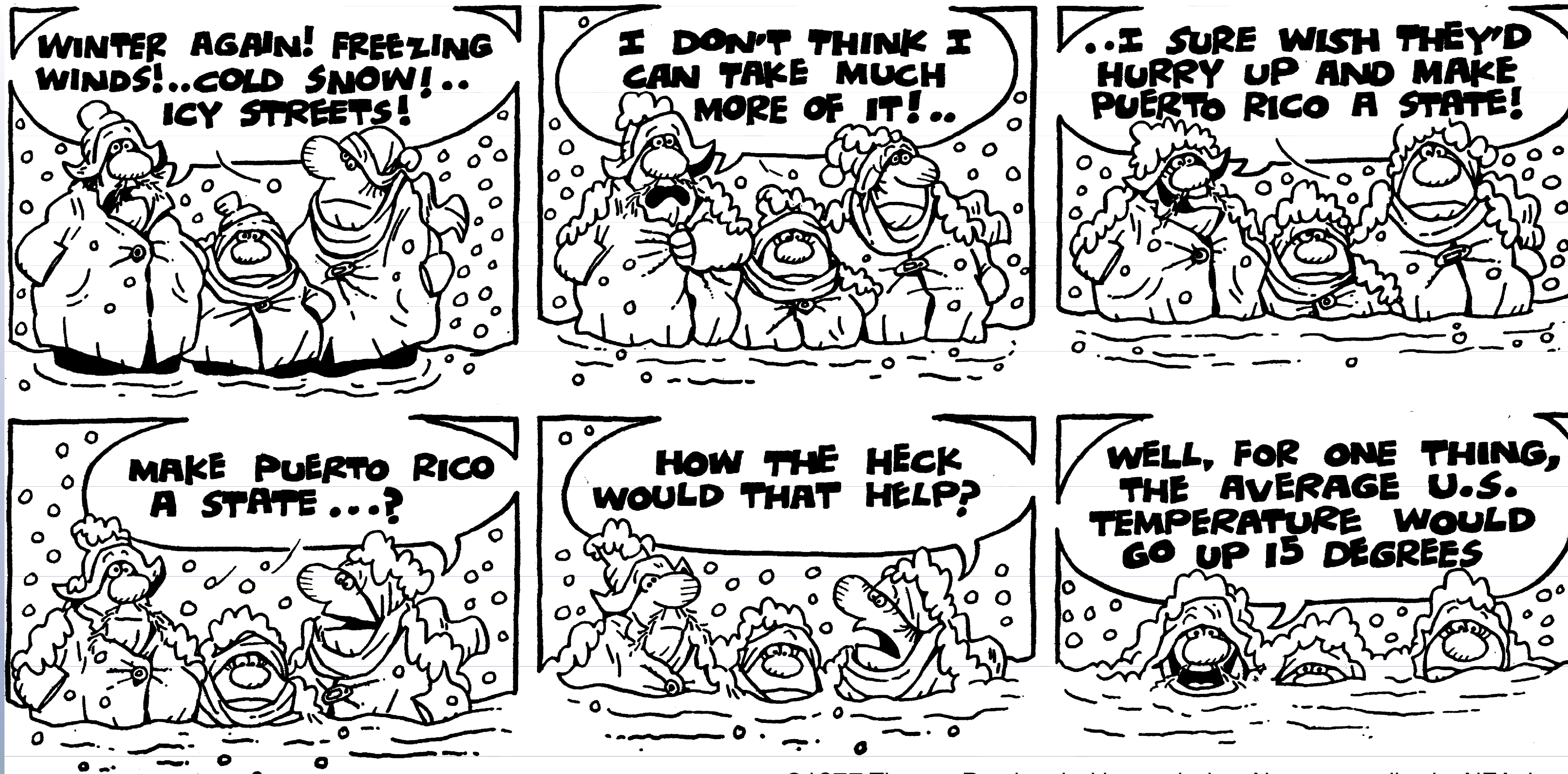
Grin and Bear It — By Lichty



"There are 2 and 3-10ths people to see you, Chief."

Moments and Percentiles (cont.)

Frank and Ernest



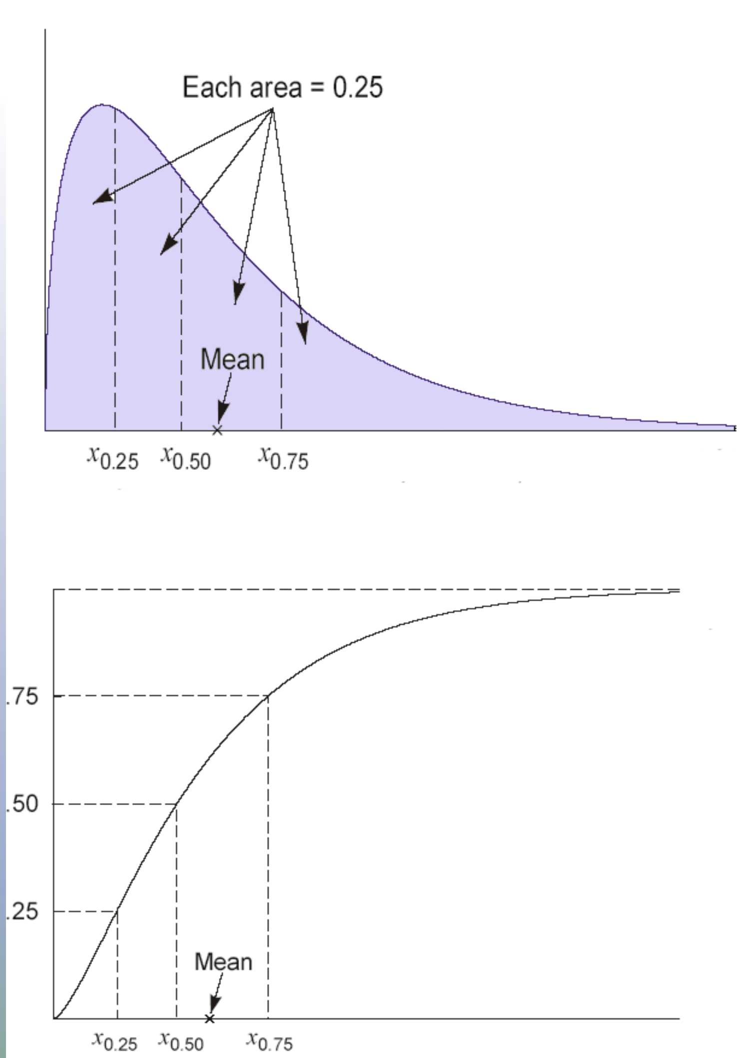
Moments and Percentiles (cont.)

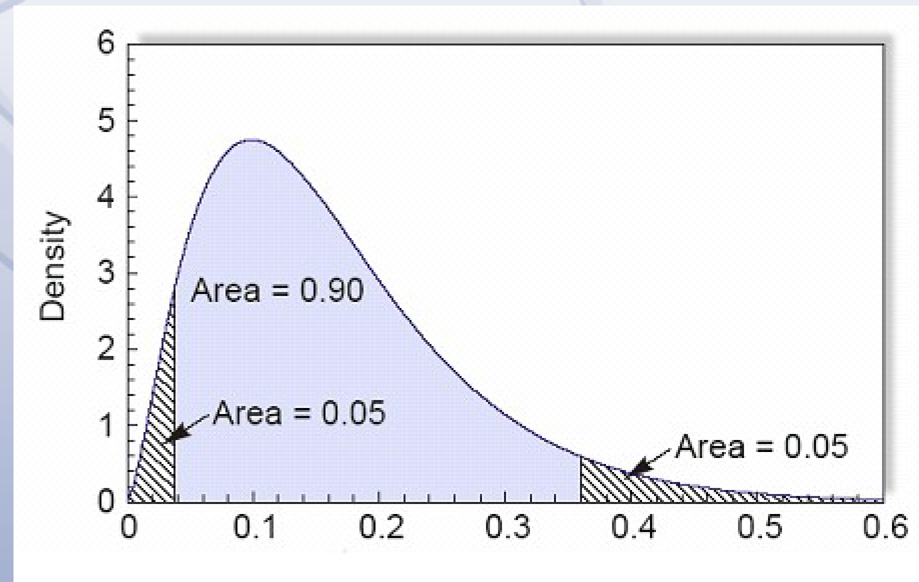
- $X \sim$ **Poisson**(λt)
  - $E(X) = \lambda t$
  - $Var(X) = \lambda t$
- $X \sim$ **binomial**(n, p)
  - $E(X) = np$
  - $Var(X) = np(1 - p)$
- $T \sim$ **exponential**(λ)
  - $E(T) = 1/\lambda$
  - $Var(T) = 1/\lambda^2$
- $X \sim$ **normal**(μ, σ^2)
  - $E(X) = \mu$
  - $Var(X) = \sigma^2$

Moments and Percentiles (cont.)

- *The 95th* ***percentile****, denoted* $x_{0.95}$*, is the value such that* $F(x_{0.95}) = 0.95$
  - *Primarily used for continuous distributions*
- *Similar definition for any number from 0 to 100 percent*
- *Special cases common in PRA include*
  - ***Median*** *= 50th percentile*
  - *1st and 3rd* ***quartiles*** *= 25th and 75th percentiles*
  - ***Upper bound*** *=* 95^{th}
    - *Should properly be called* $95^{\%}$ *upper bound*
  - ***Lower bound*** *=* 5^{th}
    - *Should properly be called the 5% lower bound*

Moments and Percentiles (cont.)





Moments and **Percentiles (cont.)**

- *Alternative language*
  - *The q **quantile** is the 100q percentile*
- *If a distribution is **positively skewed** (longer tail on the right), then mean is greater than median*
  - *E[x] > 50th*
  - *Also, the mode (highest point on the pdf) is less than the 50th percentile for positively skewed distributions*
    - *Mode < Median < Mean*

Relations among Distributions

- *In a* **Poisson** *process with event frequency λ*
  - *Time to first event and between events is* **exponential**(λ)
    - *See HOPE, page A-17*
- *If n is large, p is small, and np is moderate, then*
  - **Binomial**(*n, p*) *distribution is approximately* **Poisson**(μ)
    - *In this case,* $\mu = np$
    - *See HOPE, page A-14*
- *Normal approximations (p. A-16, plus pictures)*
  - **Binomial**(*n, p*) *is approximately* ***normal*** *as n becomes large and p is fixed*
  - **Poisson**(μ) *is approximately* ***normal*** *as* μ *becomes large*