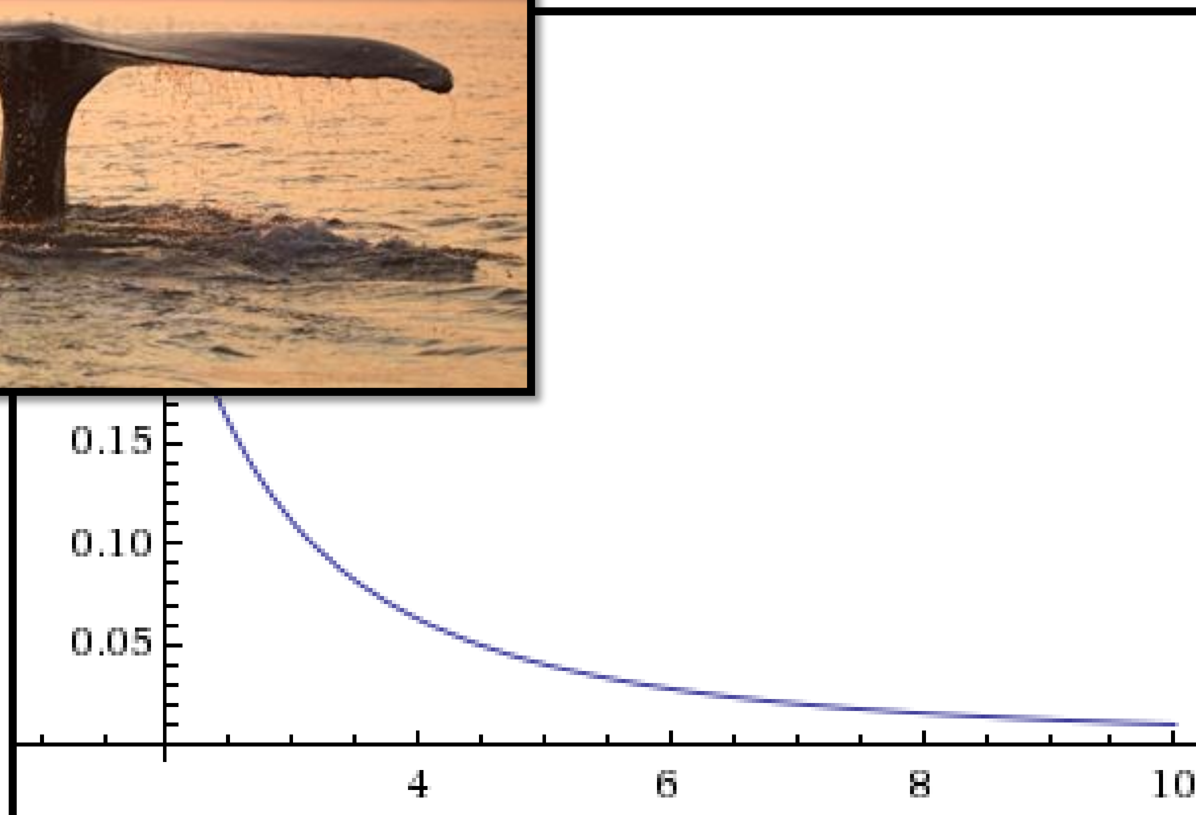
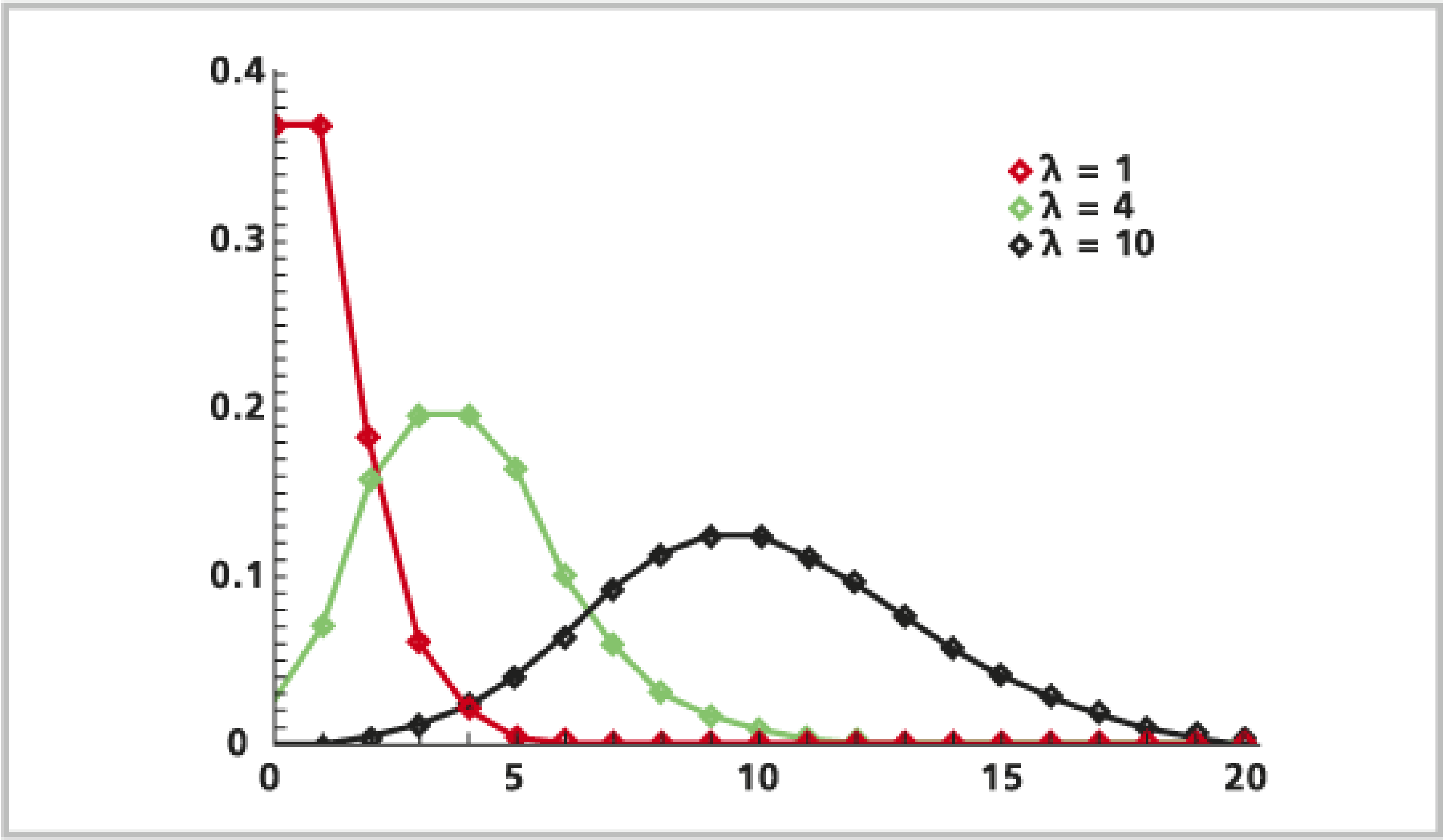
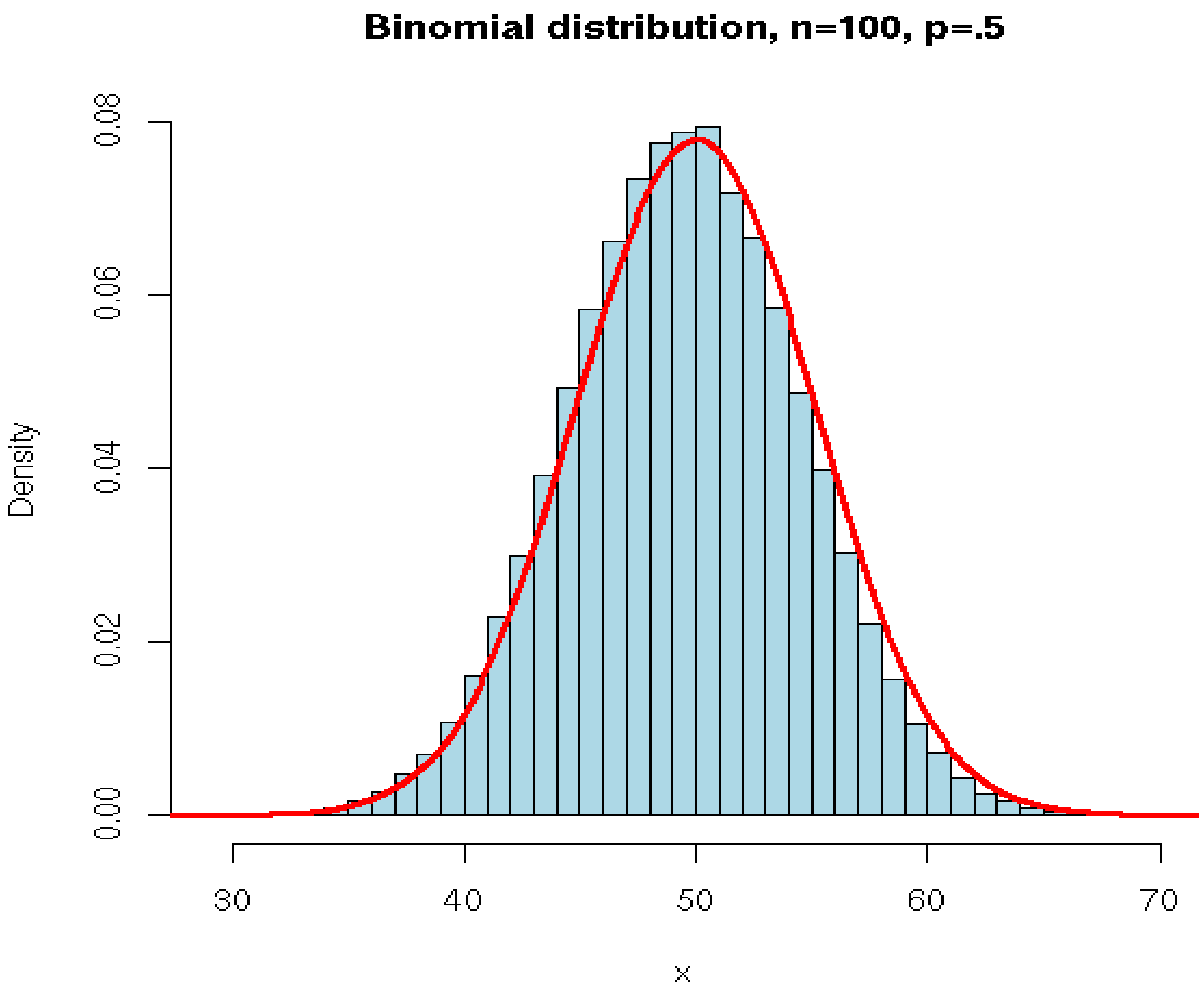
# tail bounds



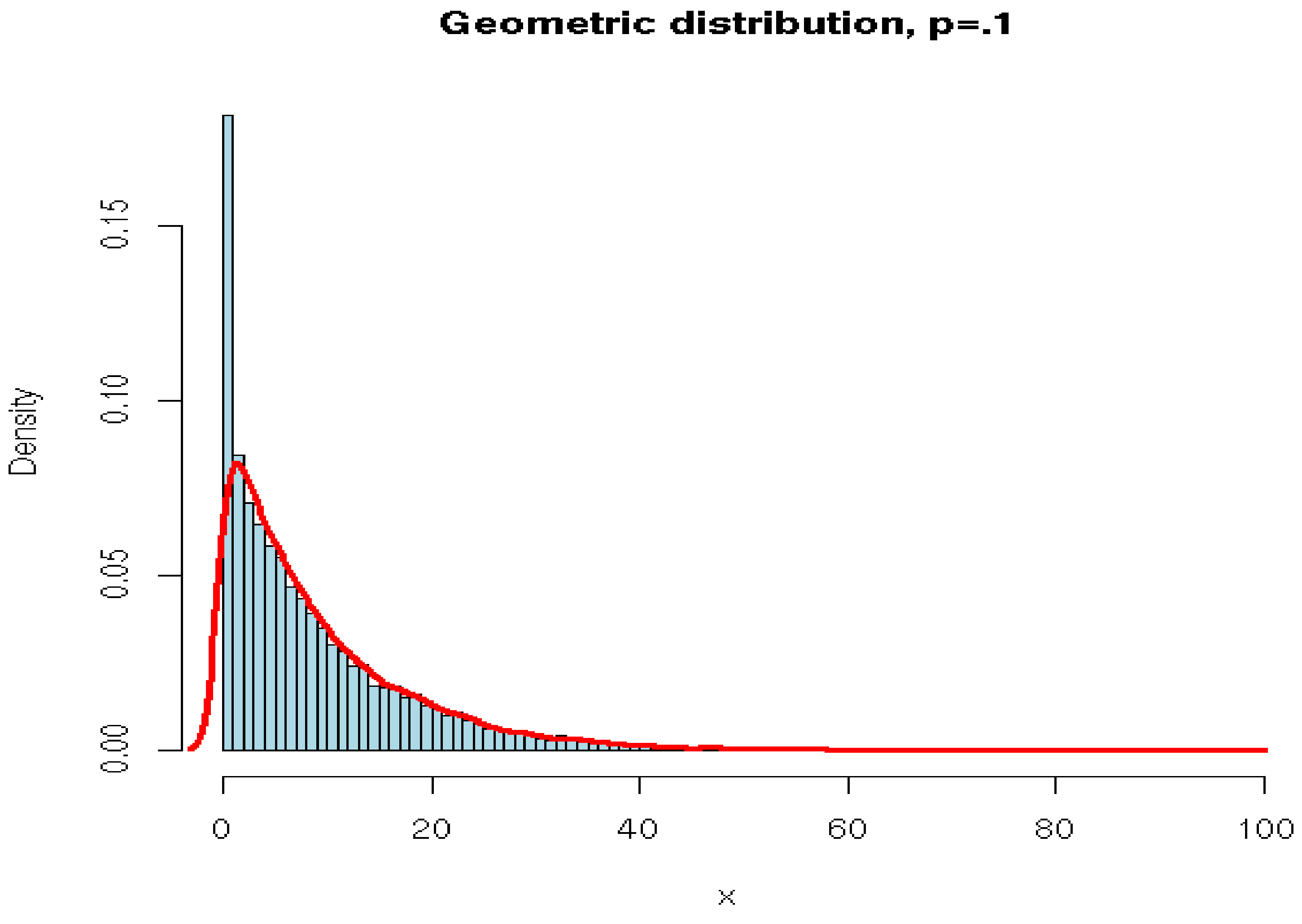
- If we have some random variable X, then the tail of X is the part of the PMF which is "far" from its expectation.



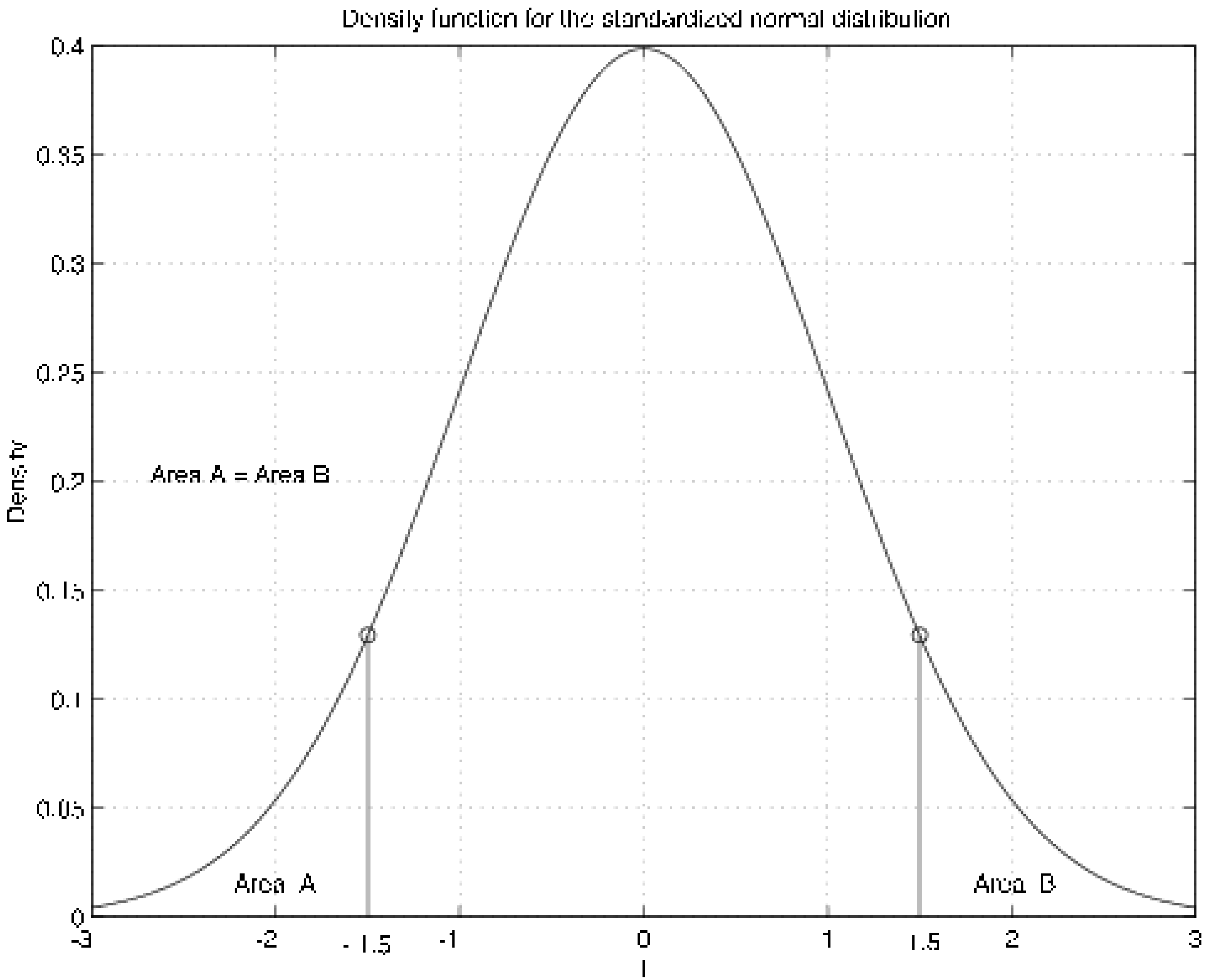
# binomial tail



# geometric tail



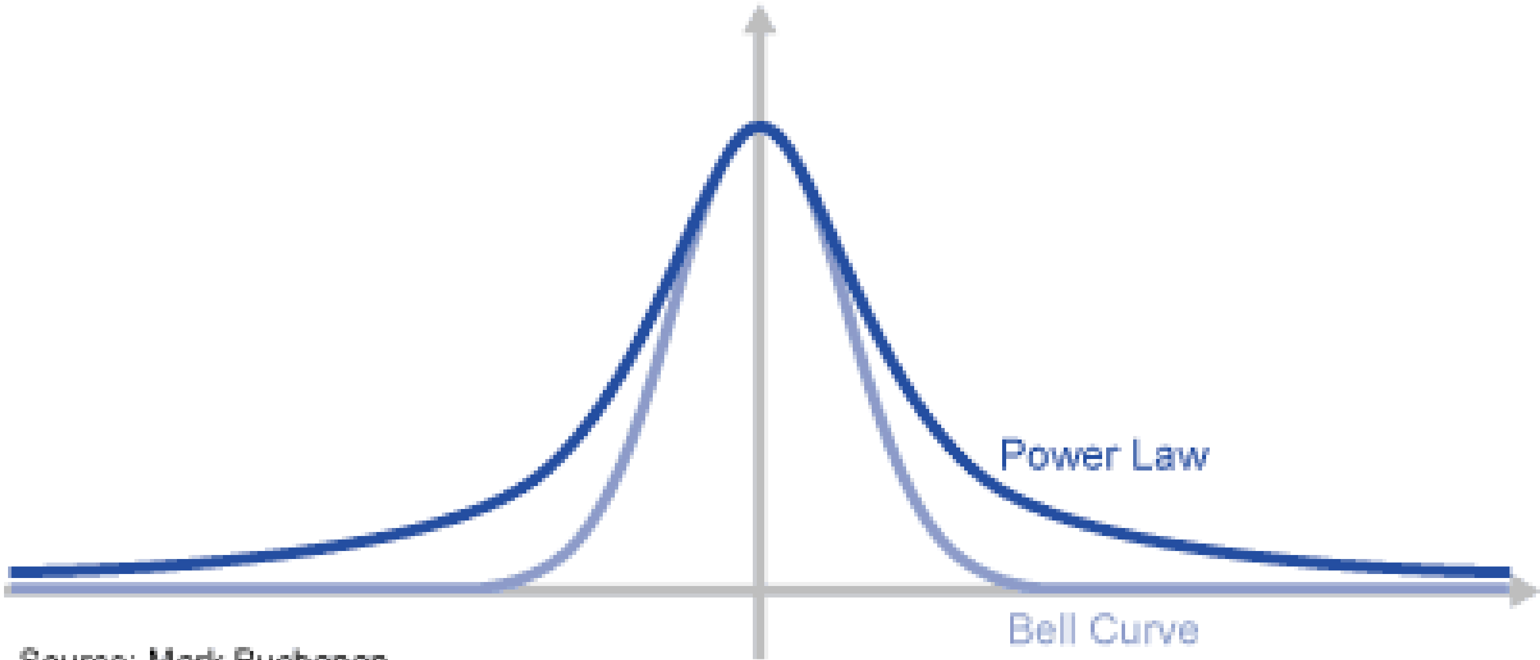
# normal tail



Density function for the standardized normal distribution
Density
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Area A = Area B
Area A
Area B
-3
-2
-1.5
-1
0
1
1.5
2
3

# heavy-tailed distribution

Exhibit 1:
**The Bell Curve vs. the Power Law: The Importance of "Fat Tails"**




Source: Mark Buchanan

- We want to bound the probability that a random variable X is large. For example:

$$P(X > \alpha) < \frac{1}{\alpha^3}$$

$$P(X > E[X] + t) < e^{-t}$$

$$P(|X - E[X]| > t) < \frac{1}{\sqrt{t}}$$

# applications of tail bounds

- We know that randomized quicksort runs in O(n log n) **expected** time. But what's the probability that it takes more than $10 \cdot n \cdot \log(n)$ steps? More than $n^{1.5}$ steps?

- Question on HW5: We know the expected advertising cost is $1500/day, but what's the probability we go over budget?

- I only expect 10,000 homeowners to default on their mortgages. What's the probability that 1,000,000 homeowners default?

# Markov's inequality

- In general, an arbitrary random variable could have very bad behavior.

- Suppose we know that X is always non-negative.

- **Theorem:** If X is a non-negative random variable, then for every $\alpha > 0$, we have

$$P(X \geq \alpha) \leq \frac{E[X]}{\alpha}$$

# Markov's inequality

- **Theorem:** If X is a non-negative random variable, then for every $\alpha > 0$, we have

$$P(X \geq \alpha) \leq \frac{E[X]}{\alpha}$$

- For example, if X = money spent on advertising in a day
- E[X] = 1500
- Then, by Markov's inequality,

$$P(X \geq 6000) \leq \frac{1500}{6000} = 0.25$$

# Markov's inequality

- **Theorem:** If X is a non-negative random variable, then for every $\alpha > 0$, we have

$$P(X \geq \alpha) \leq \frac{E[X]}{\alpha}$$

Proof:

$$E[X] = \sum_{x:p(x)>0} x \cdot p(x)$$

$$\geq \alpha \cdot P(X \geq \alpha)$$

- **Theorem:** every

Proof:

# Chebyshev's inequality

- If we know **more** about a random variable, we can often use that information to get **better** tail bounds.

- Suppose we know the variance of X.

- **Theorem:** If X is an arbitrary random variable with $\mu$ = E[X], then, for any $\alpha > 0$,

$$P(|X - \mu| > \alpha) \leq \frac{\mathrm{Var}[X]}{\alpha^2}$$

- **Theorem:** If Y is an arbitrary random variable with $\mu$ = E[Y], then, for any $\alpha > 0$,

$$P(|Y-\mu| \geq \alpha) \leq \frac{\mathrm{Var}[Y]}{\alpha^2}$$

Proof: $\text{Let } X = (Y-\mu)^2$

X is non-negative, so we can apply Markov's inequality:

$$P(|Y-\mu| \geq \alpha) = P(X \geq \alpha^2)$$

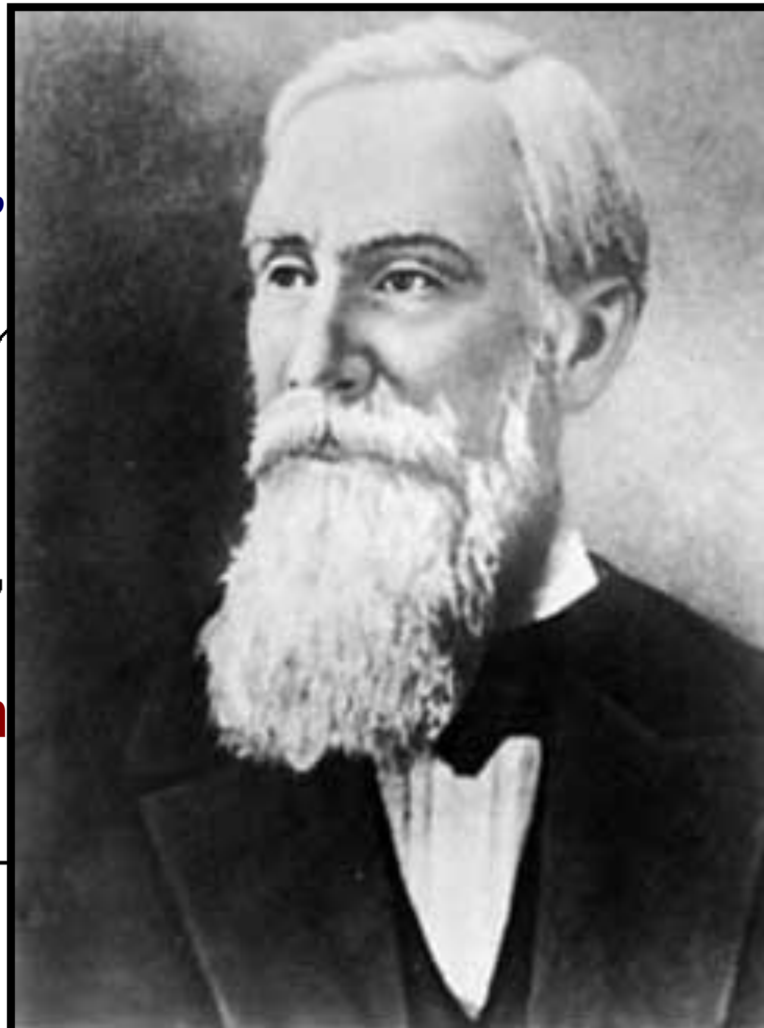
$$\leq \frac{E[X]}{\alpha^2} = \frac{\mathrm{Var}[Y]}{\alpha^2}$$

# Chebyshev's inequality

- **Theorem:** If Y with $\mu$ = E[Y], then,

$$P(|Y \quad \frac{\text{Var}[Y]}{\alpha^2}$$

Proof: Let



X is n arkov's inequality:

$$P(|Y - \quad \geq \alpha^2)$$

$$\leq \frac{\quad}{\alpha^2} = \frac{\text{Var}[Y]}{\alpha^2}$$

# Chebyshev's inequality

$$P(|Y - \mu| \geq \alpha) \leq \frac{\mathrm{Var}[Y]}{\alpha^2}$$

- Y = money spent on advertisiting in a day
- E[Y] = 1500
- Var[Y] = $500^2$ (i.e. SD[Y] = 500)

$$P(Y \geq 6000) = P(|Y - \mu| \geq 4500)$$

$$\leq \frac{500^2}{4500^2} = \frac{1}{81} \approx 0.012$$

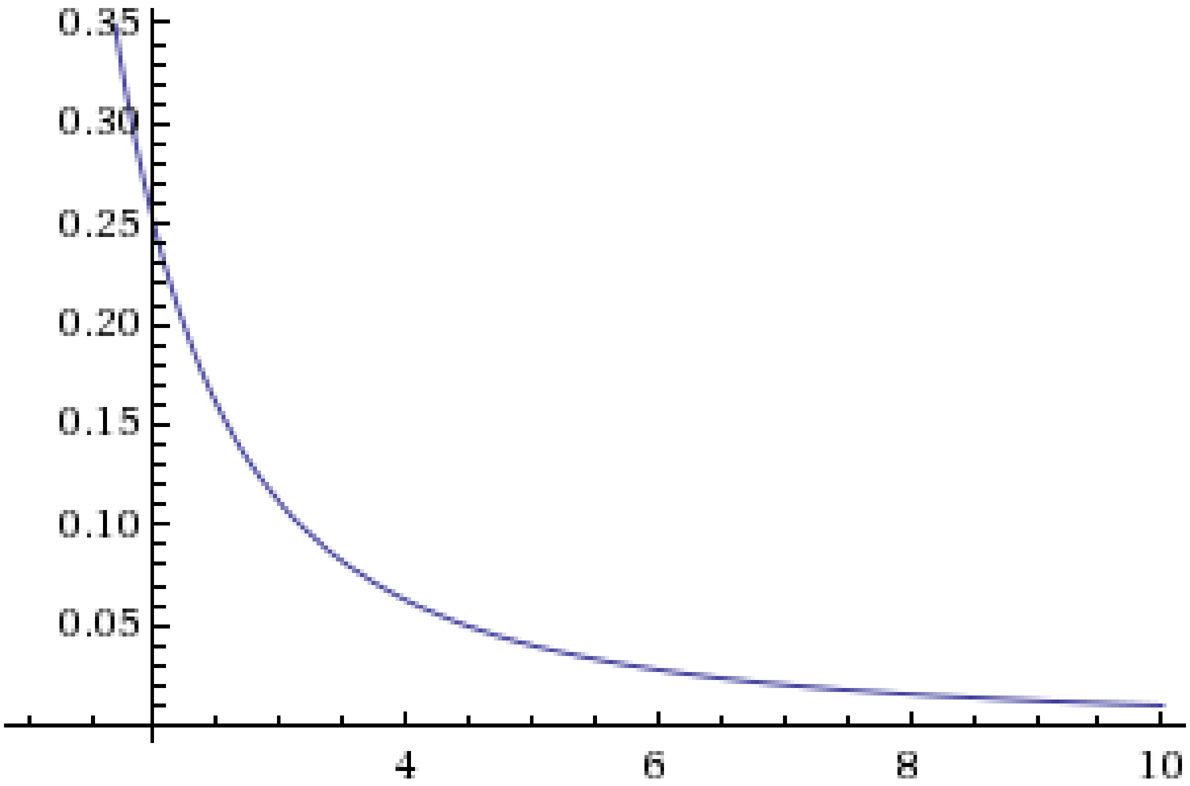
# Chebyshev's inequality

- **Theorem:** If Y is an arbitrary random variable with $\mu$ = E[Y], then, for any $\alpha > 0$,

$$P(|Y - \mu| \geq \alpha) \leq \frac{\mathrm{Var}[Y]}{\alpha^2}$$

$$\sigma = SD[Y] = \sqrt{\mathrm{Var}[Y]}$$

$$P(|Y - \mu| \geq t\sigma) \leq \frac{\sigma^2}{t^2\sigma^2} = \frac{1}{t^2}$$




$$P(|Y - \mu| \geq t\sigma) \leq \frac{1}{t^2}$$

- Y = money spent on advertisiting in a day
- E[Y] = 1500
- $Y \sim \mathrm{Bin}(15000, 0.1)$

- Poisson approximation: $Y \sim \mathrm{Poi}(1500)$
- Rough computer calculation:

$$P(Y \geq 6000) \leq 10^{-1600}$$

# Chernoff bounds

- Suppose X ~ Bin(n,p)
- $\mu$ = E[X] = pn

- Chernoff bound:

For any $\delta$ with $0 < \delta < 1$,

$$P(X > (1+\delta)\mu) \leq e^{-\frac{\delta^2 \mu}{2}}$$

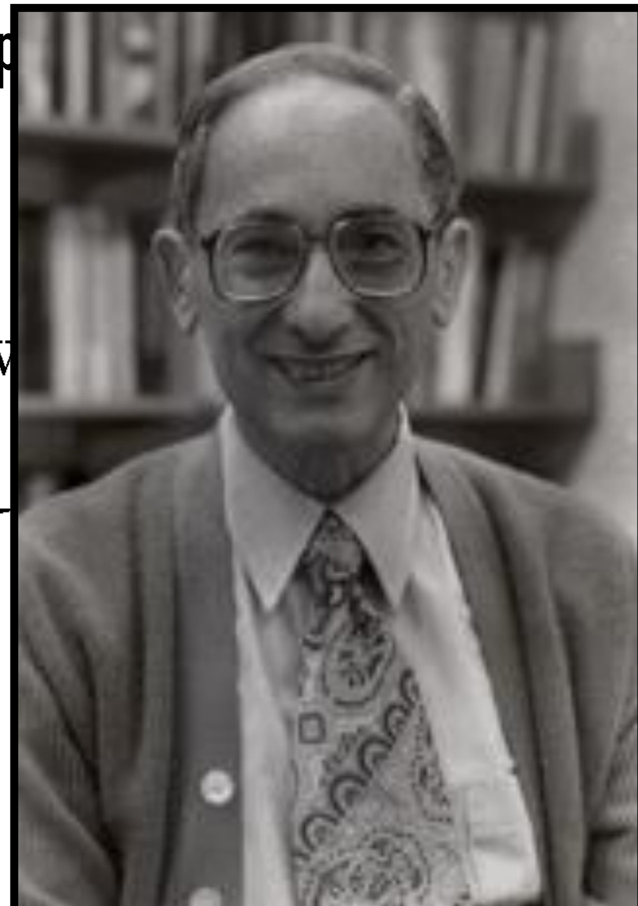
$$P(X < (1-\delta)\mu) \leq e^{-\frac{\delta^2 \mu}{3}}$$

# Chernoff bounds

- Suppose X ~ Bin(n,p)
- $\mu$ = E[X] = p

- Chernoff bound:

  For any $\delta$ w



$$P(X \quad e^{-\frac{\delta^2 \mu}{2}}$$

$$P(X \quad e^{-\frac{\delta^2 \mu}{3}}$$

# router buffers

- **Model:** 100,000 computers. Each independently sends a packet with probability p = 0.01 every second. The router processes its buffer every second as well.
- How many packets does the buffer need to hold so that the router never drops a packet?

  100,000
- What if we want to drop a packet with probability at most $10^{-6}$, every hour?

  1148
- What if we want to drop a packet with probability at most $10^{-40}$, every year?

  1300

- $X \sim \text{Bin}(100{,}000, 0.01)$
- $\mu = E[X] = 1000$

- By Chernoff bound,

$$P(X > (1+\delta)\mu) \leq e^{-\frac{\delta^2 \mu}{2}}$$

- Choose this probability to be at most: $(1/60)(1/60)(10^{-6})$

$$\delta \approx 0.148$$

- Choose this probability to be at most: $(1/60)(1/60)(1/24)(1/365)(10^{-40})$

$$\delta \approx 0.331$$