Algorithms
Professor John Reif

ALG 2.0
Probability Theory

(a) Random Variables: Binomial and Geometric

(b) Useful Probabilistic Bounds and Inequalities

Main Reading Selections:
CLR, Chapter 6
Auxillary Reading Selections:
BB, Chapter 8
Handout: "Probability Theory Refresher"

A **probability measure** *(Prob)* is a mapping from a set of events to the reals such that

(1) For any event A

$0 \leq Prob(A) \leq 1$

(2) Prob (*all* possible events) = 1

(3) If A,B are mutually exclusive events, then

$Prob(A \cup B) = Prob(A) + Prob(B)$

Conditional Probability

define $\text{Prob}(A|B) = \dfrac{\text{Prob}(A \wedge B)}{\text{Prob}(B)}$

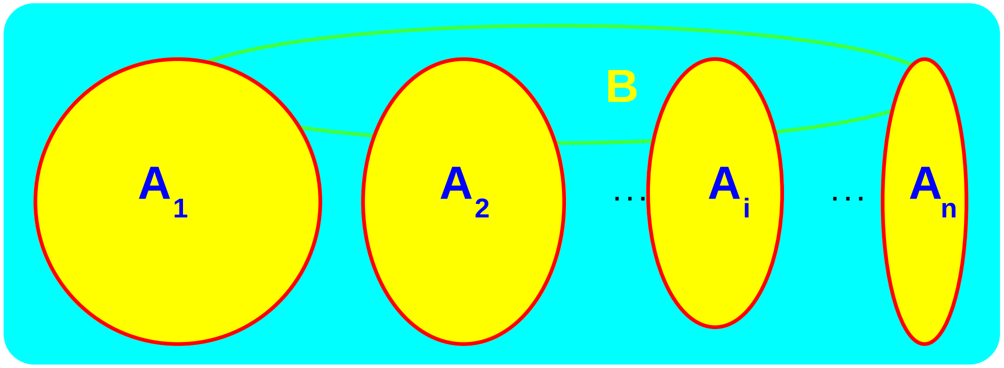
for Prob(B) > 0

Bayes' Theorem

If $A_1, \ldots, A_n$ are mutually exclusive and contain all events

then $$\text{Prob}(A_i|B) = \frac{P_i}{\sum_{j=1}^{n} P_j}$$

where $P_j = \text{Prob}(B|A_j) \cdot \text{Prob}(A_j)$

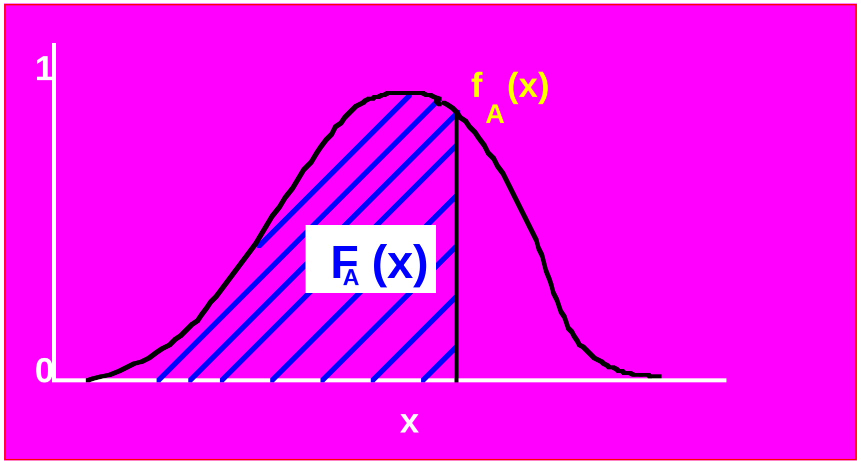


Random Variable **A**

(over real numbers)

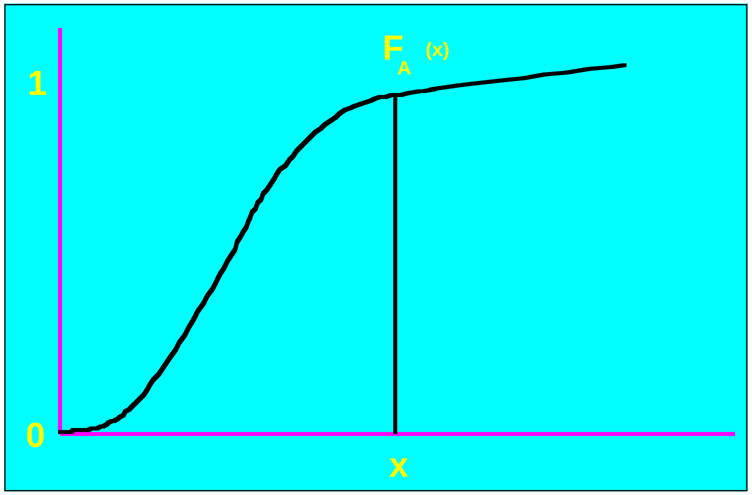
Density Function

$f_A(x)$=Prob(A=x)



prob Distribution Function

$$F_A(x) = \text{Prob}\ (A \le x) = \int_{-\infty}^{x} f_A(x)\ dx$$



If for Random Variables A,B

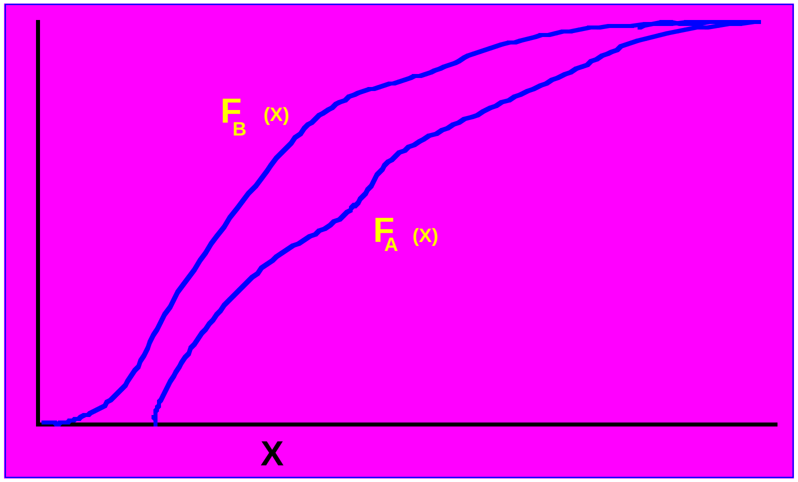
$$\forall x \quad F_A(x) \le F_B(x)$$

then

"A upper bounds B"

and

"B lower bounds A"



$F_A(x) = \text{Prob } (A \le x)$

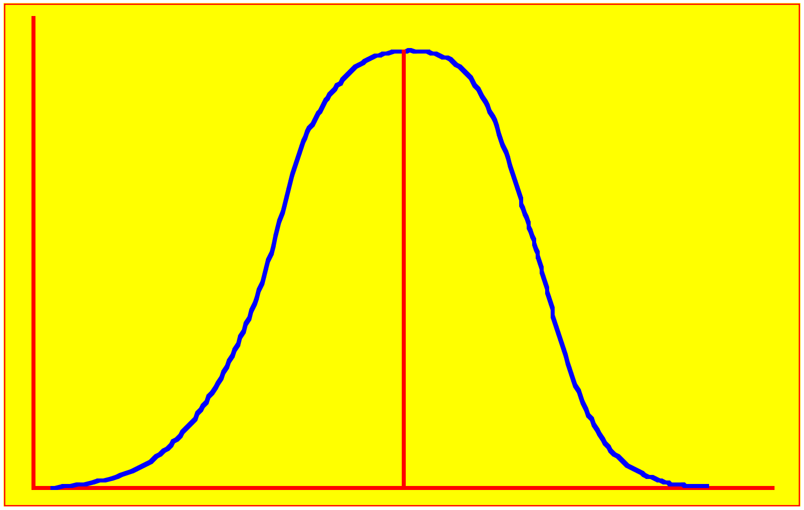
$F_B(x) = \text{Prob } (B \le x)$

Expectation of Random Variable A

$$E(A) = \bar{A} = \int_{-\infty}^{\infty} x\, f_A(x)\, dx$$

$\bar{A}$ **is also called** ***"average of A"***

and ***"mean of A"*** $= \mu_A$

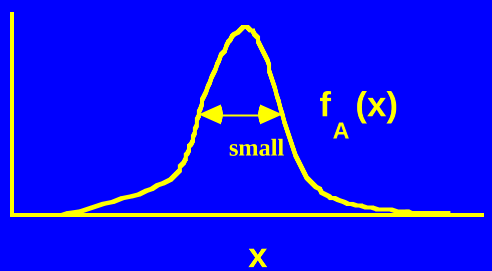


Variance of Random Variable A

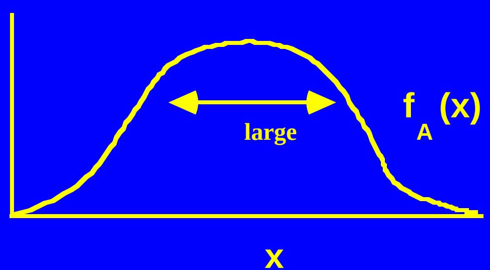
$$\sigma_A^2 = \overline{(A-\bar{A})^2} = \overline{A^2} - (\bar{A})^2$$

where *2nd moment* $\overline{A^2} = \int_{-\infty}^{\infty} x^2 f_A(x)\, dx$

example **small variance**



example **large variance**



n'th Moments of Random Variable A

$$\overline{A^n} = \int_{-\infty}^{\infty} x^n f_A(x)\, dx$$

moment generating function

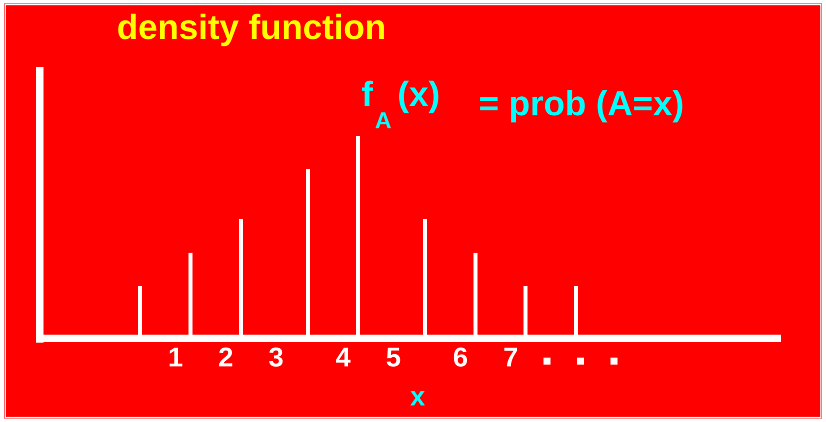
$$M_A(s) = \int_{-\infty}^{\infty} e^{sx} f_A(x)\, dx$$

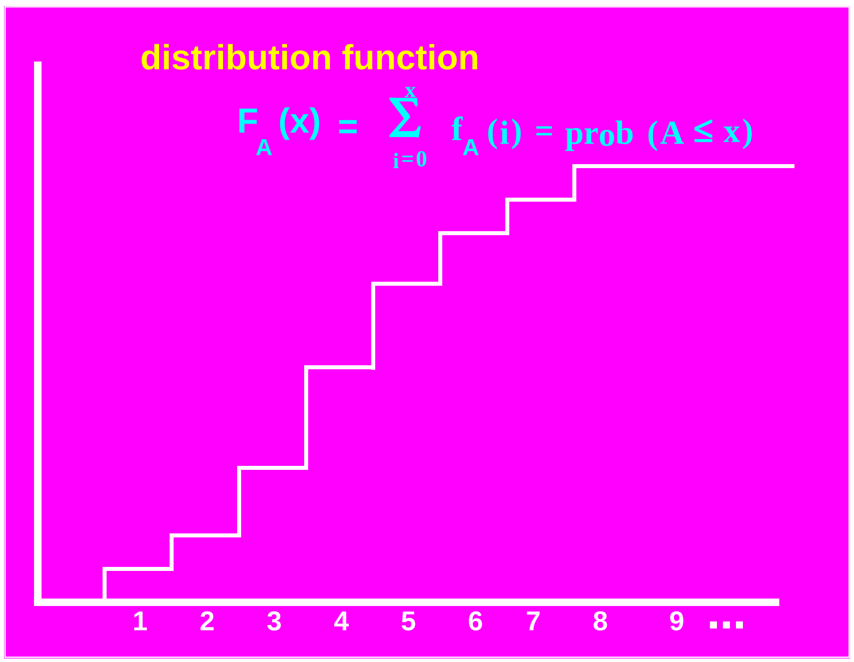
$$= E\left(e^{sA}\right)$$

note **s is a formal parameter**

$$\overline{A^n} = \left[\frac{d^n M_A(s)}{ds^n}\right]_{s=0}$$

Discrete Random Variable A





Discrete Random Variable A over nonnegative integers

expectation $E(A) = \overline{A} = \sum_{x=0}^{\infty} x \; f_A(x)$

n'th moment $\overline{A^n} = \sum_{x=0}^{\infty} x^n \; f_A(x)$

probability generating function

$$G_A(z) = \sum_{x=0}^{\infty} z^x \; f_A(x) = E\left(z^A\right)$$

1st derivative $G'_A(1) = \overline{A}$

2nd derivative $G''_A(1) = \overline{A^2} - \overline{A}$

variance $\sigma_A^2 = G''_A(1) + G'_A(1) - (G'_A(1))^2$

A,B ***independent if***

$$\mathrm{Prob}(A \wedge B) = \mathrm{Prob}(A) \cdot \mathrm{Prob}(B)$$

equivalent definition of independence

$$f_{A \wedge B}(x) = f_A(x) \cdot f_B(x)$$

$$M_{A \wedge B}(s) = M_A(s) \cdot M_B(s)$$

$$G_{A \wedge B}(z) = G_A(z) \cdot G_B(z)$$

If $A_1, \ldots, A_n$ ***independent*** **with same distribution**

$$f_{A_1}(x) = f_{A_i}(x) \quad \text{for } i = 1, \ldots, n$$

Then if $B = A_1 \wedge A_2 \wedge \ldots \wedge A_n$

$$f_B(x) = \left(f_{A_1}(x)\right)^n$$

$$M_B(s) = \left(M_{A_1}(s)\right)^n, \quad G_B(z) = \left(G_{A_1}(z)\right)^n$$

Combinatorics

$$n! = n \cdot (n-1) \cdots 2 \cdot 1$$

= number of permutations of n objects

Stirling's formula

$$n! = f(n)\ (1+o(1))$$

where $f(n) = n^n e^{-n} \sqrt{2\pi n}$

note

tighter bound

$$f(n)\, e^{\frac{1}{(12n+1)}} < n! < f(n)\, e^{\frac{1}{12n}}$$

$\frac{n!}{(n-k)!}$ **= number of permutations of n objects taken k at a time**

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

= number of (unordered)

combinations of n objects

taken k at a time

Bounds (due to Erdos & Spencer, p. 18)

$$\binom{n}{k} \sim \frac{n^k e^{-\frac{k^2}{2n} - \frac{k^3}{6n^2}}}{k!} (1-o(1))$$

for $k = o\left(n^{\frac{3}{4}}\right)$

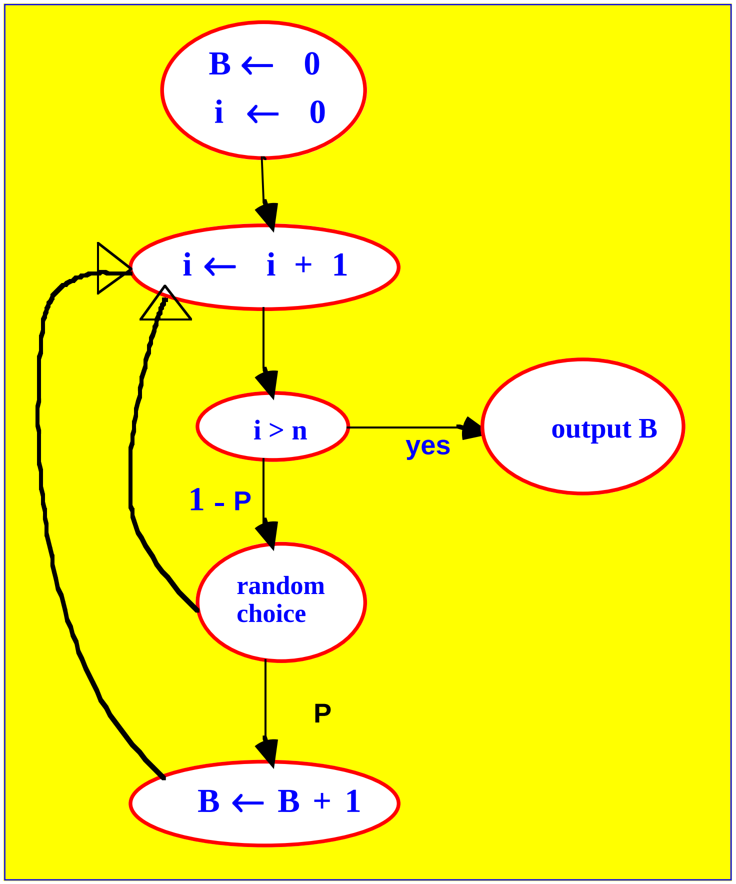
Bernoulli Variable

A_i is 1 with prob P and 0 with prob 1-P

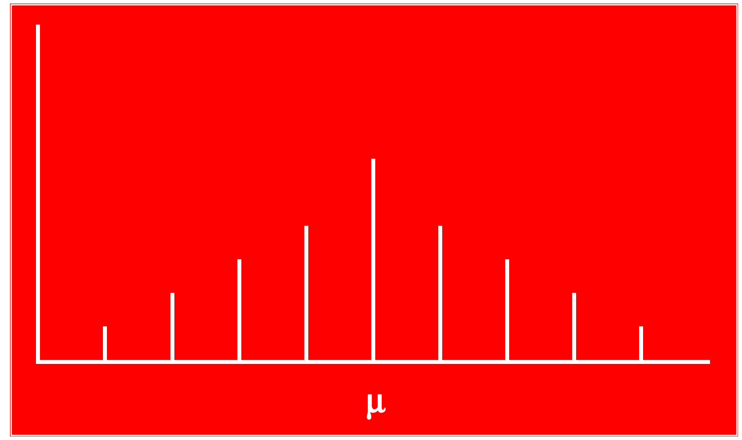
Binomial Variable

B is sum of n independent Bernoulli variables A_i each with some probability p

```
procedure   BINOMIAL with parameters n,p
  begin       B ← 0
              for i=1 to n do
                  with probability P    do B ← B+1
              output B
  end
```



B is Binomial Variable with parameters n,p



mean $\mu = n \cdot p$

variance $\sigma^2 = np\,(1-p)$

density fn $= \text{Prob}(B=x) = \binom{n}{x} p^x \, (1-p)^{n-x}$

distribution fn $= \text{Prob}(B \le x) = \sum_{k=0}^{x} \binom{n}{k} p^k \, (1-p)^{n-k}$

GENERATING FUNCTION

$$G(z) = (1-p+pz)^n = \sum_{k=0}^{n} z^k \binom{n}{k} p^k (1-p)^{n-k}$$

interesting fact

$$\text{Prob}(B=\mu) = \Omega\left(\frac{1}{\sqrt{n}}\right)$$

Poisson Trial

A_i is 1 with prob P_i
and 0 with prob $1-P_i$

Suppose B' is the

sum of n independent Poisson trials

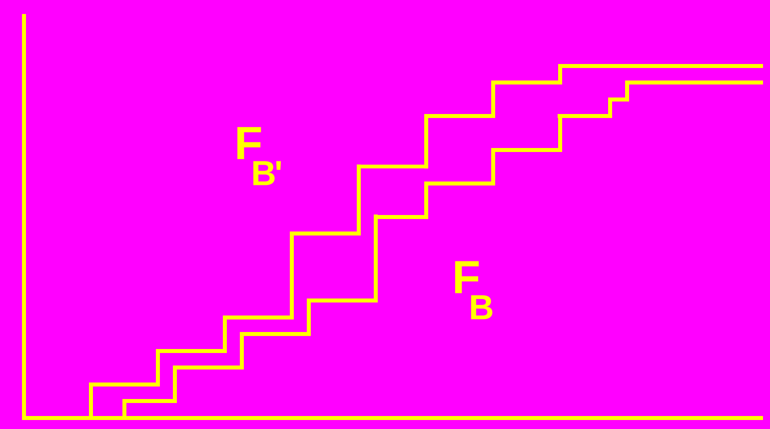
A_i with probability P_i for $i > 1,\ldots,n$

Hoeffding's Theorem

B' is *upper bound*

by a Binomial Variable B

parameters n, p where $p = \frac{\sum_{i=1}^{n} P_i}{n}$



Geometric Variable

V parameter p

$\forall x \geq 0 \quad Prob(V=x) = p(1-p)^x$

procedure

GEOMETRIC parameter p

begin $V \leftarrow 0$

loop: with probability p

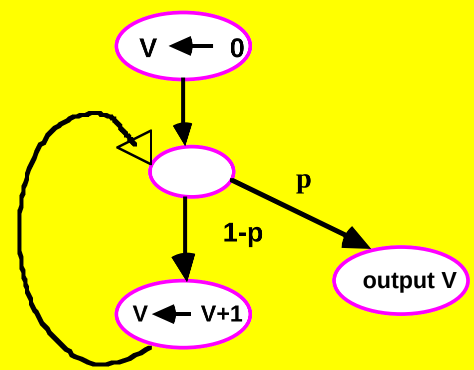
goto exit

$V \leftarrow V+1$

goto loop

exit: *output* V

mean $\mu = \frac{1-p}{p}$



GENERATING FUNCTION

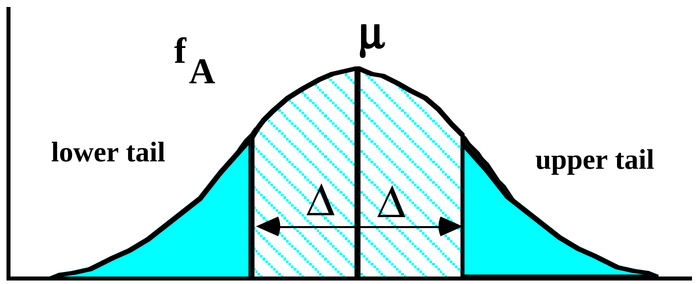
$$G(Z) = \sum_{k=0}^{\infty} Z^k \,(p(1-p)^k) = \frac{p}{1-(1-p)Z}$$

Probabilistic Inequalities

for Random Variable A

mean $\mu = \overline{A}$

variance $\sigma^2 = \overline{A^2} - (\overline{A})^2$



Markov Inequality **(uses only *mean*)**

$$\text{Prob } (A \geq x) \leq \frac{\mu}{x}$$

Chebychev Inequality **(uses *mean* and *variance*)**

$$\text{Prob } (|A-\mu| \geq \Delta) \leq \frac{\sigma^2}{\Delta^2}$$

example

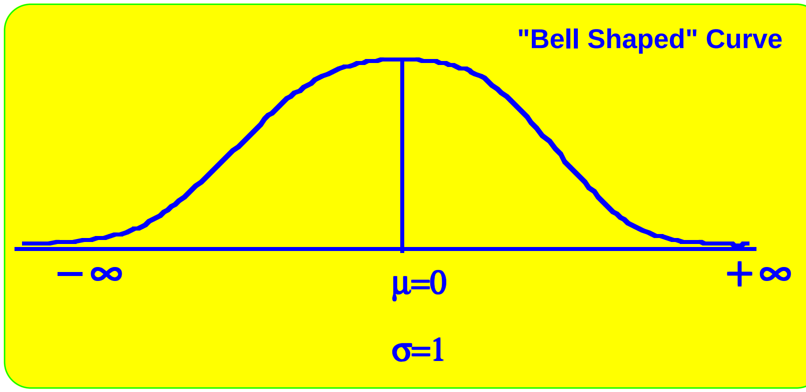
If B is a Binomial **with parameters n,p**

$$\text{Then Prob } (B \geq x) \leq \frac{np}{x}$$

$$\text{Prob } (|B-np| \geq \Delta) \leq \frac{np\,(1-p)}{\Delta^2}$$

Gaussian Density function

$$\Psi(x) = \frac{1}{\sqrt{2\Pi}}\, e^{-\frac{x^2}{2}}$$



Normal Distribution

$$\Phi(X) = \int_{-\infty}^{x} \Psi(Y)\, dY$$

Bounds

$\forall x > 0$

(Feller, p. 175)

$$\Psi(x)\left(\frac{1}{x} - \frac{1}{x^3}\right) \leq 1 - \Phi(x) \leq \frac{\Psi(x)}{x}$$

$\forall x \; \varepsilon \; [0,1]$

$$\frac{x}{\sqrt{2\Pi e}} = x\,\Psi(1) \leq \Phi(x) - \frac{1}{2} \leq x\,\Psi(0) = \frac{x}{\sqrt{2\Pi}}$$

Let S_n be the

sum of n independently distributed variables

$A_1, \ldots, A_n$

each with mean $\frac{\mu}{n}$ and variance $\frac{\sigma^2}{n}$

So S_n has mean μ and variance σ^2

Strong Law of Large Numbers

The probability density function of

$$T_n = \frac{(S_n - \mu)}{\sigma} \textit{ limits} \text{ as } n \to \infty$$

to *normal distribution* $\Phi(x)$

Hence Prob

$(|S_n - \mu| \leq \sigma x) \to \Phi(x)$ as $n \to \infty$

so Prob

$(|S_n - \mu| \geq \sigma x) \to 2(1-\Phi(x))$

$\leq 2\,\Psi(x)/x$

(since $1-\Phi(x) \leq \Psi(x)/x$)

Chernoff Bound of Random Variable A

(uses *all* moments)

Prob $(A \geq x) \leq e^{-sx} M_A(s)$ for $s \geq 0$

$= e^{\gamma(s)-sx}$ where $\gamma(s) = \ln(M_A(s))$

$\leq e^{\gamma(s) - s\gamma'(s)}$

(by setting $x = \gamma'(s)$

1st derivative minimizes bounds)

need *moment generating function*

Chernoff Bound of *Discrete* Random Variable A

$$\text{Prob}(A \geq x) \leq z^{-x} G_A(z) \quad \text{for } z \geq 1$$

choose $z = z_0$ to *minimize* above bound

need *Probability Generating* function

$$G_A(z) = \sum_{x \geq 0} z^x f_A(x) = E\left(z^A\right)$$

Chernoff Bounds for *Binomial B* with parameters n,p

Above mean $x \geq \mu$

$$\text{Prob}(B \geq x)$$

$$\leq \left(\frac{n-\mu}{n-x}\right)^{n-x} \left(\frac{\mu}{x}\right)^{x}$$

$$\leq e^{x-\mu} \left(\frac{\mu}{x}\right)^{x} \quad \text{since } \left(1 - \frac{1}{x}\right)^{x} < e^{-1}$$

$$\leq e^{-x-\mu} \quad \text{for } x \geq \mu e^{2}$$

Below Mean $x \le \mu$

Prob (B ≤ x)

$$\le \left(\frac{n-\mu}{n-x}\right)^{n-x} \left(\frac{\mu}{x}\right)^{x}$$

Anguin-Valiant's Bounds

for

Binomial B

with parameters n,p

Just above mean

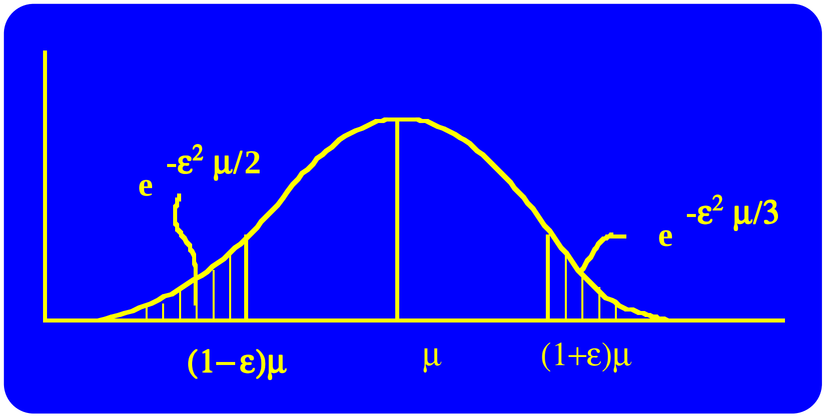
$\mu = np$ for $0 < \varepsilon < 1$

$$\text{Prob}\ (B \ge (1+\varepsilon)\mu) \le e^{-\varepsilon^2 \frac{\mu}{2}}$$

Just below mean

μ for $0 < \varepsilon < 1$

$$\text{Prob}\ (B \le \lfloor (1-\varepsilon)\mu \rfloor) \le e^{-\varepsilon^2 \mu/3}$$



⇒ **tails are bounded by Normal distributions**

Binomial Variable B
with Parameters p, N
and expectation μ = pN

By Chernoff
Bound for p < 1/2

$$\text{Prob}\ (B \geq \frac{N}{2}) < 2^N p^{\frac{n}{2}}$$

Raghavan-Spencer bound	For any $\partial > 0$

$$\text{Prob}\ (B \geq (1+\partial)\mu) \leq \left(\frac{e^{\partial}}{(1+\partial)^{(1+\partial)}} \right)^{\mu}$$

in FOCS'86.