Mathematics of Computer Science. - M.I.T. opencourseware

Abstract:
Large deviation means there are lot of gap between the event point and expected value. So it is hard to figure out the meaningful boundary. However with Markov and Chebyshev and Chernoff's theory, but also there are some limit conditions, we could tellthe probability's bound.

For the English:
develop, say, trivially, might, state of the art, minuscule

Lec 24. LARGE DEVIATION:

- Markov's THM:

If R is a non-negative rv, then ∀x > 0. Pr(R>=x) =< Ex(R)/x

PF: Ex(R) = Ex(R|R>=x)*Pr(R>=x) + Ex(R|R<x)*Pr(R<x) >= 0 (non-negative)

=> Ex(R) >= xPr(R>=x)

COR: If R is non-negative rv, then ∀c > 0 Pr(R >= cEx(R)) =< 1/c

PF: set x = cEx(R) in THM (Ex(R) be eliminated by denominator)

EX: R = weight of random person (weight is never negative)

Suppose Ex(R) = 100

Pr(R >= 200) =< 100/200 = 1/2

Chinese appetizer != hat check problem : get bound

1/n 1/n!

what if R is small upper bound

COR: If R =< u for some u ∈ |R, then ∀x < u

Pr(R =< x) =< u-Ex(R) / u-x

PF: Pr(R =< x) = R(u-R >= u-x)

Pr(u-R >= u-x) =< Ex(u-R)/u-x == u-Ex(R)/u-x

EX: R = score of random student

max score = 100 = u // suppose Ex(R) = 75

Pr(R =< 50) = 100-75/100-50 = 1/2

If know variance => could got better bounds

- Chebyshev's THM

analog of Markov's Thm based on variance

: ∀x > 0 & any rv R, Pr(|R-Ex(R)| >= x) =< Var(R)/x^2

PF: Pr(|R-Ex(R)| >= x) = Pr((R-Ex(R))^2 >= x^2) =< Ex((R-Ex(R))^2)/x^2 = Var(R)

deviation square = non-negative apply Markov thm

COR: Pr(|R-Ex(R)| >= c$(R)) =< Var(R)/c^2<u>$(R)^2</u> = 1/c^2

std dev =Var(R)

EX: R = IQ of random person

Assume R >= 0, Ex(R) = 100, $(R) = 15

Pr(R >= 250) Markov: =< 2/5 in thm
Chebyshev: Pr(R-100 >= 150) = Pr(R-Ex(R) >= 10$(R)) =< Pr(|R-Ex(R)| >= 10$(10)) = 1/100
which is much better bound
not normalized => not always symmetric

THM: For any rv R
Pr(R – Ex(R) >= c$(R)) =< 1/C^2+1 // Pr(R - Ex(R) =< -c$(R)) =< 1/c^2+1

- Chernoff Bound
THM: Let T1 ~ Tn be any mut indep rv's s/t ∀j 0 =< Tj =< 1
Let T = ∑Tj, then for any c > 1, Pr(T >= cEx(T)) =< e^-ZEx(T) - exponentially small
Where Z = cln(c) + 1 – c > 0
EX: Suppose Ex(T) = 100, c = 2 => Z = 2ln(2) + 1 – 2 > 0.38
Pr(T) >= 2Ex(T) =< e^-38 (-0.38 * 100) - exponentially small
EX: 10 million people play pick 4
Pr(win) = 1/10000 Ex(# winners) = 1000   Pr(>= 2000 winners) =< e^-380 if mut indep picks // however in Markov's thm => 1/2 too rough bound
Pr(>= 1100 winners)   c = 1.1   z = 1.1ln(1.1) + 1 – 1.1 >= 0.0048
=< e^-4.8 < 1/100

PF: Chernoff for case Tj ∈ {0, 1}:
Pr(T >= cEx(T)) = Pr(c^T >= c^cEx(T)) =< Ex(c^T)/c^cEx(T) by Markov
T = T1 + T2 ... Tn => C^T = C^T1*C^T2...C^Tn   (∏)
=> Ex(C^T) = Ex(∏C^Tj) = ∏Ex(C^Tj) by product rule & mut indep
Ex(C^Tj) = C1Pr(Tj=1) + C0Pr(Tj=0) = CPr(Tj=1) + 1 – Pr(Tj=1)
= 1 + (c-1)Pr(Tj) = 1 + (c-1)Ex(Tj)   fact : 1 + x =< e^x
=< e^(c-1)Ex(Tj) => Ex(C^T) =< ∏e^(c-1)Ex(Tj) = e^(c-1)Ex(∑Tj) - linearity of exp
= e^(c-1)Ex(T)   and then came back to Markov
=< Ex(C^T)/C^cEx(T) = e^(c-1)Ex(T)/C^cEx(T) = e^-(cln(c) + c –1)Ex(T) = e^-ZEx(T)

(A>B) = Pr(f(A) > f(B)) =< Ex(f(A))/f(B)

- Load balancer
N jobs B1, B2 ~ Bn // M servers S1 ~ Sn // N = 100,000, M = 10
Bj takes Lj times (0 =< Lj =< 1)
Let = L = ∑Lj
Assume L = 25,000 L/M = 25,000/10 = 2500
: Practical solution is assign job randomly
PF: Let Rij be load on server Si from job Bj
Rij = { Lj if Bj assigned to Si – prob = 1/m || 0 otherwise – prob = 1 – 1/m }
Let Ri be =< ∑Rij = load on Si   Ex(Ri) = ∑Ex(Rij)
deviation = 서버간의 편차
Pr(Ri >= C^L/M) =< e^-2L/M z= cln(c) + 1 – c = ∑Lj/M = L/M is optimal