**Probability Distributions**

* Random variable – value is determined by chance
  + E.g. rolling a die:
  + X 1 2 3 4 5 6
  + P(X) 1/6 1/6 1/6 1/6 1/6 1/6
  + X is a discrete variable
* For a probability distribution:
  + ∑ P(X) = 1
  + 0 ≤ P(X) ≤ 1
* **Mean** of a probability distribution
  + μ = ∑ (X ⋅ P(X))
* **Variance** of a probability distribution
  + σ2 = ∑ (X2 ⋅ P(X)) – μ2
* **Standard deviation**
  + σ = sqrt(σ2)
* **Expectation**
  + E(X) = μ
  + For a game
    - E(X) > 0 → in favour of the player e.g. win money
    - E(X) < 0 → not in favour of the player e.g. lose money
    - E(X) = 0 → fair game
* **Binomial distribution & binomial experiment**
  + Each trial has 2 outcomes (success & failure)
  + Each outcome has the same probability for every trial
    - p = P(S); q = 1 – p = P(F); both are constant
  + There are a fixed number of trials (n)
  + All trials are independent
  + X = # of successes in n trials
    - i.e. 0 ≤ X ≤ n
  + P(X=k) = C(n, k) ⋅ pk ⋅ qn-k
    - P(0) + … P(n) = qn + npqn-1 + … + npn-1q + pn = (p + q)n = 1 (Binomial expansion)
  + E.g. tossing a coin 3 times
    - P(X=0) = C(3, 0) ⋅ (1/2)3 ⋅ (1/2)0 = 1/8
    - P(X=1) = C(3, 1) ⋅ (1/2)2 ⋅ (1/2) = 3/8
    - P(X=2) = 3/8
    - P(X=3) = 1/8
    - E(X) = μ = 0(1/8) + 1(3/8) + 2(3/8) + 3(1/8) = 12/8 = 1.5 (heads out of 3 tosses)
    - Var(X) = 0 + 12 ⋅ 3/8 + 22 ⋅ 3/8 + 32 ⋅ 1/8 – 1.52 = 0.75
    - σ(X) = sqrt(0.75)
  + **If X ~ Bin(n, p)** – X follows a binomial distribution
    - Then E(X) = np and Var(X) = np(1 – p)
* **Poisson distribution**
  + Useful for when n is large and p is small
  + X has a Poisson distribution with parameter λ > 0 if its probability mass function (pmf) is given by
    - P(X; λ) = P(X=x) = where x = 0, 1, 2, 3…
    - λ is a rate
    - P(X=0) + P(X=1) + … = e^-λ ⋅ ∑(n = 0 → ∞) λn/n! = e^-λ ⋅ e^λ = 1 (Maclauren’s)
  + **Mean/expectation**: μ = λ = np
  + **Variance**: σ2 = λ
  + Ex: λ = 3 website visit per hour
    - P(at most 2 visit)

= P(X=0) + P(X=1) + P(X=2)

= e-3 + 3 ⋅ e-3 + 9/2 ⋅ e-3 = 17/2 ⋅ e-3

* + - P(exactly 7 visits in 5 hours) – λ = 5\*3 = 15

= P(X=7) = 157/7! ⋅ e-15

* **Poisson approximation for binomial distributions**
  + For n >> (very large) and p << (very small) such that np ~= λ, we can approximate:
    - **Bin(n, p) ~= Poiss(λ)**
  + Ex: P(any given page has at least 1 error) = 0.005; total of 400 pages in the book
    - Poisson approx. λ = 400(0.005) = 2
    - P(exactly one page with errors) = 2e-2
    - P(at most 3 pages with errors) = P(X=0) + P(X=1) + P(X=2) + P(X=3)
      * = e-2 + 2e-2 + 22/2! ⋅ e-2 + 23/3! ⋅ e-2
* **Poisson process**
  + Let λ = average rate (per unit time)
  + Let X = # of events that occur in t units of time
  + Then X has a Poisson distribution i.e. X ~ Poiss(λ ⋅ t)
  + Pmf of x = f(x) =
  + E(x) = var(x) = λt
  + Ex: calls to 911 ~ a Poisson process with λ = 3 calls/minute
    - P(6 calls in the next 2.5 min)

= f(6)|t=2.5 = (3\*2.5)6/6! ⋅ e-3(2.5) = 0.137

* + - P(2 calls in the first min (A) | 6 calls in the next 2.5 min (B))

= P(A ∩ B)/P(B) ← must make A & B independent so we can use P(A)P(B)

= P(2 calls in 1 min & 4 calls in (the following) 1.5 min)/P(B)

= f(2)|t=1 ⋅ f(4)|t=1.5 / 0.137 = 0.311

* **Cumulative distribution function (CDF)**
  + CDF of a discrete random variable X is given by:
    - F(x) = P(X ≤ x) for any x
  + E.g. P(X=0) = 0.1, P(X=1) = 0.2, P(X=3) = 0.4, P(X=5) = 0.3
    - F(0) = 0.1
    - F(1) = 0.3
    - F(3) = 0.7
    - F(5) = 1
* **Geometric distribution**
  + **Bernoulli trial** – trials w/ 2 independent outcomes and w/ constant probability of success (p) for all trials
  + Let X = # of trials until (and including) the first success, then X follows a geometric distribution
  + i.e. X ~ Geom(p) and its pmf is given by f(x) = P(X=x) = p(1 – p)x – 1
  + E(X) = 1/p
  + Var(X) = (1 – p)/x2
  + E.g. p = 1/6 chance of winning every cup
    - Chance of winning on the 10th cup = P(X=10) = (1/6)(5/6)9 = 0.0323
  + **Memory-less property of geometric random variable**
    - Let x be a geometric random variable and t1, t2 ∈ R
    - Then P(X ≥ t1 + t2 | X ≥ t1) = P(X ≥ t2)
    - E.g. chance of winning after 30 cups given having failed 20 cups = chance of winning after 10 cups
* **Chebyshev’s Theorem** – revisited
  + Let X be a r. v. w/ E(X) = Var(X) = σ2; then for any ε > 0
  + P(|X – E(X)| ≤ ε) ≥ Var(X)/ε2 for any ε > 0
  + P(|X – E(X)| < ε) = 1 – P(|X – E(X)| ≥ e) ≥ 1 – Var(X)/ε2
  + Let ε = kσ where k > 0, standard deviation σ > 0; then
  + P(|X – E(X)| < kσ) ≥ 1 – 1/k2
  + Note that |X – E(X)| < kσ ⇔ -kσ < X – E(X) < kσ
  + E.g. # of students who miss class is r. v. X with mean = 15 and σ = 2
    - P(9 < X < 21) = P(-6 < X – 15 < 6)
    - = P(|X – 15| ≤ 6) ← kσ = 6, k = 3
    - ≥ 1 – 1/32
  + E.g. X ~ Bin(5, ½)
    - E(X) = np = 2.5
    - Var(X) = np(1 – p) = 1.25
    - P(|X – E(X)| < 2σ) = P(|X – 2.5| < 2sqrt(1.25)) ≥ 1/22