Probability and Probability Distributions

1 Probability

1.1 Introduction: Lotteries as losing bets ("voluntary taxes")

Sucker Bet: wager in which expected return is significantly lower than the wager.

EXPECTED RETURN = winning prize * P(win) − cost of wager * P(lose).

1.2 Preliminaries -- see Zar chapter 5

Counting Possible Outcomes; Permutations; Combinations

Prob. of an Event: rel. frequency = frequency of event / total number all events

Range: [0,1]

Adding Probabilities

Mutually exclusive events (e.g., heads/tails): P(A or B) = P(A) + P(B)

Not mutually exclusive (i.e., intersecting): P(A or B) = P(A) + P(B) − P(A and B)

Multiplying Probabilities

P(A and B) = P(A)*P(B)

1.3 Probability density function (pdf): P(X = x), i.e., prob that variate has value x

For continuous distribution [since P(X = x) = 0], usually $P(x_a \le X \le x_b)$

1.4 Cumulative distribution function (cdf): $P(X \le x)$, i.e., prob that variable ≤ x

2 Probability Distributions

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

2.1 Probability distribution summary

Distribution	pdf	Mean	Std.Dev.	CV
Uniform	$f(x) = \frac{1}{B-A}$	(A+B)/2	$\sqrt{\frac{(B-A)^2}{12}}$	$\frac{(B-A)}{\sqrt{3}(B+A)}$
Binomial	$P(x,p,N) = \binom{N}{x}(p)^{xk}(1-p)^{(n-x)}$	Np	$\sqrt{Np(1-p)}$	$\sqrt{\frac{(1-p)}{np}}$
Normal	$f(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$	μ	σ	σ/μ
t		0	$\sqrt{\frac{\nu}{(\nu-2)}}$	undefined
F		$\nu_2/(\nu_2-2)$		
Chi-square	$f(x) = \frac{e^{-x/2}x^{\frac{\nu}{2}-1}}{2^{\nu/2}\Gamma(\nu/2)}$	ν	$\sqrt{2\nu}$	$\sqrt{\frac{2}{\nu}}$
Poisson	$p(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$	λ	$\sqrt{\lambda}$	$\frac{1}{\sqrt{\lambda}}$
Lognormal	$f(x) = \frac{e^{-\left[\frac{\ln\{(x-\theta)/m\}^2}{2\sigma^2}\right]}}{(x-\theta)\sigma\sqrt{2\pi}}$	$e^{0.5\sigma^2}$	$\sqrt{e^{\sigma^2}(e^{\sigma^2}-1)}$	$\sqrt{e^{\sigma^2}-1}$

$$\binom{N}{x} = \frac{N!}{x!(N-x)!}$$

$$\Gamma_x(a) = \int t^{a-1} e^{-t} dt$$

2.2 Continuous Distributions

2.2.1 Uniform Distribution

pdf: $f(x) = \frac{1}{B-A}$ for $A \le x \le B$

i.e., standard uniform pdf: $f(x) = 1$ for $0 \le x \le 1$

graph: horizontal line at P = 1

cdf: $f(x) = x$ for $0 \le x \le 1$

mean: $(A + B)/2$ SD: $\sqrt{\frac{(B-A)^2}{12}}$ CV: $\frac{B-A}{\sqrt{3}(B+A)}$

2.2.2 Normal Distribution: from Central Limit Theorem

pdf: $f(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$

pdf of standard normal ($\mu = 0$, $\sigma = 1$): $f(x) = \frac{e^{-x^2/2}}{\sigma\sqrt{2\pi}}$

mean: μ SD: σ CV: μ/σ

2.2.3 *t* Distribution

leptokurtic (heavy tails) relative to Normal distribution

mean: μ SD: $\sqrt{\frac{\nu}{\nu-2}}$

2.2.4 Chi-square Distribution

sum of ν independent standard normal distributions, each squared

pdf: $f(x) = \frac{e^{-x/2} x^{\frac{\nu}{2}-1}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)}$ for $x \ge 2$, $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$

mean: ν SD: $\sqrt{2\nu}$ CV: $\sqrt{2/\nu}$

2.2.5 *F* Distribution: ratio of two Chi-square distributions

mean: $\frac{\nu_2}{(\nu_2 - 2)}$, $\nu_2 > 2$ CV: $\sqrt{\frac{2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)}}$, $\nu_2 > 4$

2.3 Discrete Distributions

2.3.1 Binomial Distribution: 2 mutually exclusive outcomes per event

probability mass function: $p(x,p,n)=\binom{n}{x}p^{x}(1-p)^{n-x}$, for x = 0, 1, 2, …, n

where $\binom{n}{x}=\dfrac{n!}{x!(n-x)!}$

cumulative probability function: $F(x,p,n)=\sum_{i=0}^{x}\binom{n}{x}p^{i}(1-p)^{n-i}$

mean: np SD: $\sqrt{np(1-p)}$ CV: $\sqrt{\dfrac{1-p}{np}}$

2.3.2 Poisson Distribution: e.g., number of events w/in time (or space) interval

probability mass function: $p(x,\lambda)=\dfrac{e^{-\lambda}\lambda^{x}}{x!}$, for x = 0, 1, 2, …

mean: λ SD: $\sqrt{\lambda}$ CV: $1/\sqrt{\lambda}$

3 Conditional Probability

Bayes Theorem: P(B|A) = P(AB) / P(A)