Discrete-type random variables

pmf (probability mass function):$p_X(u) = P\{X = u\}$
$E[g(X)] = \sum_i g(u_i)p_X(u_i)$.
$E[ag(X) + bh(X) + c] = aE[g(X)] + bE[h(X)] + c$.
The variance and standard deviation of a random variable:
$Var(X) = E[X^2] - \mu_X^2$, $E[aX + b] = aE[X] + b$, $Var(aX + b) = a^2 Var(X)$
The *standardized version* of X is the random variable $\frac{X-\mu_X}{\sigma_X}$, and $Var\left(\frac{X-\mu_X}{\sigma_X}\right) = 1$
The **conditional probability** of B given A is defined by: $P(B|A) = \begin{cases} \frac{P(AB)}{P(A)} \text{ if } P(A) > 0 \\ \text{undefined if } P(A) = 0 \end{cases}$
Event A is **independent** of event B if $P(AB) = P(A)P(B)$.
Events A,B and C are **pairwise independent** if $P(AB) = P(A)P(B), P(AC) = P(A)P(C), P(BC) = P(B)P(C)$
Events A,B and C are **independent** if they are *pairwise independent* and if $P(ABC) = P(A)P(B)P(C)$

Binomial distribution

A random variable X is said to have the **Bernoulli distribution** with parameter p, where $0 \le p \le 1$, if $p_X(1) = p$ and $p_X(0) = 1 - p$. $E[X] = p$, $Var(X) = E[X^2] - E[X]^2 = p(1 - p)$
Suppose n independent *Bernoulli trials* are conducted, each resulting in a one with probability p and a zero with probability $1 - p$. Let X denote the total number of ones occurring in the n trials. The pmf of X

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ for } 0 \le k \le n$$

$E[X] = np$ and $Var[X] = np(1 - p)$

Geometric distribution

Do Bernoulli trials untill the outcome of a trial is one. L denote the number of trials conducted.

$$p_L(k) = (1-p)^{k-1}p \text{ for } k \ge 1$$

and $P\{L > k\} = (1 - p)^k$ for $k \ge 0$.
$E[L] = \frac{1}{p}$, $Var[L] = \frac{1-p}{p^2}$

Negative binomial distribution

Let S_r denotes the number of trials required for r ones, and the last trail must be one. Let $n \ge r$, and let $k = n - r$. The event $\{S_r = n\}$ is determined by the outcomes of the ffirst n trials. The event is true iff there are $r - 1$ ones and k zeros in the first $k + r - 1$ trials, and trail n is one.

$$p(n) = \binom{n-1}{r-1} p^r (1-p)^{n-r} \text{ for } n \ge r$$

$E[S_r] = \frac{r}{p}$, $Var(S_r) = rVar(L_1) = \frac{r(1-p)}{p^2}$

Poisson disttribution

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k!} \text{ for } k \ge 0, \lambda = np. E[Y] = Var[Y] = \lambda$$

Maximum likelihood parameter estimation: For a random variable X, and that the pmf of X is p_θ, where θ is a parameter. The probability of k being the observed value for X. The *likelihood* $X = k$ is $p_\theta(k)$. The *maximum likelihood estimate* of θ for observation k, denoted by $\hat{\theta}_{ML}(k)$, is the value of θ that maximizes the likelihood, $p_\theta(k)$, with respect to θ. (Give k, find θ (or p in some Bernoulli trials) to make $p_\theta(k)$ biggest, $\hat{\theta}_{ML}(k)$ =the value of θ).

Markov's inequality: $P\{Y \ge c\} \le \frac{E[Y]}{c}$

Chebychev inequality: $P\{|X - \mu| \ge d\} \le \frac{\sigma^2}{d^2}$

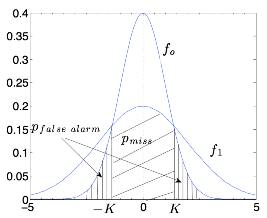
The law of total probability: $P(E_i|A) = \frac{P(AE_i)}{A} = \frac{P(A|E_i)P(E_i)}{P(A)}$ $E[X] = \sum_{j=1}^{J} E[X|E_j]P(E_j)$

confidence intervals: $P\{p \in (\hat{p} - \frac{a}{2\sqrt{n}}, \hat{p} + \frac{a}{2\sqrt{n}})\} \ge 1 - \frac{1}{a^2}$

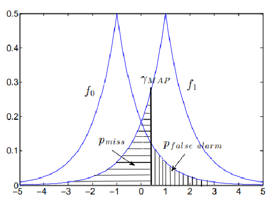
Binary hypothesis testing with discrete-type observations

	X=0	X=1	X=2	X=3
H_1	0.0	0.1	$\underline{0.3}$	$\underline{0.6}$
H_0	$\underline{0.4}$	$\underline{0.3}$	0.2	0.1

In this ML example: $p_{false-alarm} = P\{\text{declare } H_1|H_0\} = 0.2 + 0.1 = 0.3$, $p_{miss} = P\{\text{declare } H_0|H_1\} = 0.0 + 0.1 = 0.1$



$N(0, 1)$ and $N(0, 4)$ pdfs and ML threshold K.



Laplace densities centered at ± 1, and γ_{MAP} for $\frac{\pi_0}{\pi_1} = 2$.

Define the **likelihood ratio**: $\Lambda(k) = \frac{p_1(k)}{p_0(k)}$
Define a (π_0, π_1). $P(H_i, X = k) = \pi_i p_i(k)$. (For ML, $(\pi_0, \pi_1) = (0.5, 0.5)$)
An LRT with threshold τ can be written as $\Lambda(X) \begin{cases} > \tau \text{ declare } H_1 \text{ is true} \\ < \tau \text{ declare } H_0 \text{ is true} \end{cases}$
$\Lambda(X) = \tau$? It will metioned in questions(like "Asume that ties are broken in favor of H_1").
The threshold τ of LRT in MAP is $\frac{\pi_0}{\pi_1}$, $p_e = \pi_0 p_{false-alarm} + \pi_1 p_{miss}$

Union bound: ¶$(A_1 \cup A_2 \cup ... \cup A_m) \le P(A_1) + P(A_2) + ... + P(A_m)$.

Continuous-type random vairables

$F_X(c) = \int_{-\infty}^{c} f_X(u)du$
$E[X] = \int_{-\infty}^{+\infty} uf_X(u)du = \int_0^1 F_X^{-1}(u)du$, $Var(X) = E[X^2] - \mu_X^2$
$E[g(X)] = \int_{-\infty}^{\infty} g(u)f_X(u)du$, $Var(g(X)) = \int_{-\infty}^{\infty} g^2(u)f_X(u)du - \left(\int_{-\infty}^{\infty} g(u)f_X(u)du\right)^2$

Uniform distribution

Let $a \le b$. A random variable X is *uniformly distributed* over the interval $[a, b]$ if

$$f_X(u) = \begin{cases} \frac{1}{b-a} & a \le u \le b \\ 0 & else. \end{cases}, E[X] = \frac{a+b}{2}, Var[X] = \frac{(a-b)^2}{12}$$

Exponential distribution

A random variable T has the exponential distribution with parameter $\lambda > 0$ if its pdf and cdf are given by

$$F_T(t) = \begin{cases} 1 - e^{-\lambda t} & t \ge 0 \\ 0 & t < 0 \end{cases}, f_T(t) = \begin{cases} \lambda e^{\lambda t} & t \ge 0 \\ 0 & else. \end{cases}, E[T^n] = \frac{n!}{\lambda^n}, Var(T) = \frac{1}{\lambda^2}$$

Memoryless property of exponential distribution:
$P(T > s + t|T > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P\{T > t\}$, $E[T|T \ge k] = E[T] + k$

The Erlang distribution

Let T_r denote the time of the r^{th} count of a Poisson process. Thus, $T_r = U_1 + ... + U_r$.

$$f_{T_r}(t) = \frac{e^{-\lambda t}\lambda^r t^{r-1}}{(r-1)!}, P\{T_r > t\} = \sum_{k=0}^{r-1} \frac{e^{-\lambda t}(\lambda t)^k}{k!}, E[T_r] = \frac{r}{\lambda}, Var[T_r] = \frac{r}{\lambda^2}$$

The Gaussian (normal) distribution

$N(\mu, \sigma^2) \sim f(u) = \frac{e^{-\frac{(u-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$. **standard** $N(0, 1)$: $f(u) = \frac{e^{-\frac{(u)^2}{2}}}{\sqrt{2\pi}}$, $\Phi(u)$, $Q(u) = 1 - \Phi(u)$

ML parameter estimation for continuous-type variables: Derivative.

Function of a random variable

Scope the problem: Get ready.
find the CDF of Y: $F_Y(c) = P\{Y \le c\} = P\{g(X) \le c\}$
Differentiate F_Y to find its derivative, which is f_Y

Failure rate: $h(t) = \lim_{\epsilon \to 0} \frac{P\{t < T \le t + \epsilon\}}{P\{T > t\}\epsilon} = \frac{f_T(t)}{1 - F_T(t)}$, $F(t) = 1 - \exp(-\int_0^t h(s)ds)$