ECSE 6520: Estimation and Detection Theory

Commonly Used Distributions

Class Notes - 5

February 9, 2014

Contents

1 Continuous Distributions

1.1 Chi-Square Distribution

- The (central) Chi-square density with k degrees of freedom is of the form:

$$f_\theta(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2} \qquad x > 0 \tag{1}$$

  where $\theta = k$, a positive integer.

- Here $\Gamma(u)$ denotes the Gamma function,

$$\Gamma(u) = \int_0^\infty x^{u-1} e^{-x} dx, \tag{2}$$

  For n integer valued $\Gamma(n+1) = n! = n(n-1)\dots1$ and $\Gamma(n+1/2) = \frac{(2n-1)(2n-3)\dots5.3.1}{2^n}\sqrt{\pi}$.

- If $Z_i \sim N(0,1)$ are i.i.d., $i = 1\dots,n$, then $X = \sum_{i=1}^n Z_i^2$ is distributed as Chi-square with n degrees of freedom.

  Some useful properties of the Chi-square random variable are as follows:

- $E[x_n] = n$; $\mathrm{var}(x_n) = 2n$
- Asymptotic relation for large n: $x_n = \sqrt{2n}N(0,1) + n$
- x_2 an exponential r.v. with mean 2, i.e. $X = x_2$ is a non-negative r.v. with probability density $f(x) = \frac{1}{2}e^{-}x/2$.
- $\sqrt{x_2}$ is a Rayleigh distributed random variable.

1.2 Gamma Distribution

- The Gamma density function is

$$f_\theta(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \qquad x > 0 \tag{3}$$

where θ denotes the pair of parameters (λ, r), $\lambda, r > 0$.

- Let $\{Y_i\}_{i=1}^n$ be i.i.d. exponentially distributed random variables with mean $1/\lambda$, specifically Y_i has density

$$f_\lambda(y) = \lambda e^{-\lambda y} \qquad y > 0 \tag{4}$$

Then the sum $X = \sum_{i=1}^n Y_i$ has a Gamma density $f(\lambda, n)$.

Other useful properties of a Gamma distributed random variable X with parameters $theta = (\lambda, r)$ include:

- $E_\theta[X] = r/\lambda$
- $\mathrm{var}_\theta(X) = r/\lambda^2$
- The Chi-square distribution with k degrees of freedom is a special case of the Gamma distribution obtained by setting Gamma parameters as follows: $\lambda = 1/2$ and $r = k/2$.

1.3 Non-central Chi-Square Distribution

- The sum of squares of independent Gaussian r.v.s with unit variances but non-zero means is called a non-central Chi-square r.v.
- Specifically, if $Z_i \sim N(\mu_i, 1)$ are independent, $i = 1, \ldots, n$, then $X = \sum_{i=1}^n Z_i^2$ is distributed as non-central Chi-square with n degrees of freedom and non-centrality parameter $\delta = \sum_{i=1}^n \mu_i^2$.
- In our shorthand we write

$$\sum_{i=1}^{n} [N(0,1) + \mu_i]^2 = \sum_{i=1}^{n} [N(\mu_i, 1)]^2 = x_{n,\delta} \tag{5}$$

The non-central Chi-square density has no simple expression of closed form. There are some useful asymptotic relations, however:

- $E[x_{n,\delta}] = n + \delta$, $\text{var}(x_{n,\delta}) = 2(n + 2\delta)$

- $\sqrt{x_{2,\mu_1^2+\mu_2^2}}$ is a Rician r.v.

1.4 Chi-square Mixture

- The distribution of the sum of squares of independent Gaussian r.v.s with zero mean but different variances is not closed form either.

- However, many statisticians have studied and tabulated the distribution of a weighted sum of squares of i.i.d. standard Gaussian r.v.s $Z_1, \ldots, Z_n$, $Z_i \sim N(0, 1)$.

- Specifically, the following has a (central) Chi-square mixture with n degrees of freedom and mixture parameter $c = [c_1, \ldots, c_n]^T$, $c_i \geq 0$:

$$\sum_{i=1}^{n} \frac{ci}{\sum_j c_j} Z_i^2 = x_{n,c} \tag{6}$$

An asymptotic relation of interest to us will be:

- $E[x_{n,c}] = 1,$, $\text{var}(x_{n,c}) = 2 \sum_{i=1}^{N} \left(\frac{ci}{\sum_j c_i} \right)^2$

- Furthermore, there is an obvious a special case where the Chi-square mixture reduces to a scaled (central) Chi-square: $x_{n,c_1} = \frac{1}{n} x_n$ for any $c \neq 0$.

1.5 Student-t distribution

- For $Z \sim N(0, 1)$ and $Y \sim x_n$ independent r.v.s the ratio $X = Z/\sqrt{Y/n}$ is called a Student-t r.v. with n degrees of freedom, denoted T_n.

- In shorthand notation:

$$\frac{N(0, 1)}{\sqrt{x_n/n}} = T_n \tag{7}$$

- The density of T_n is the Student-t density with n degrees of freedom and has the form

$$f_\theta(x) = \frac{\Gamma([n+1]/2)}{\Gamma(n/2)} \frac{1}{\sqrt{n\pi}} \frac{1}{(1+x^2/n)^{(n+1)/2}} \tag{8}$$

where $\theta = n$ is a positive integer.

Properties of interest are:

- $E[T_n] = 0 (n > 1)$, $\text{var}(T_n) = \frac{n}{n-2} (n > 2)$

- Asymptotic relation for large n: $T_n \approx N(0; 1)$.

- For $n = 1$ the mean of T_n does not exist and for $n \leq 2$ its variance is infinite.

1.6 Fischer-F

- For $U \sim x_m$ and $V \sim x_n$ independent r.v.s the ratio $X = (U/m)/(V/n)$ is called a Fisher-F r.v. with m, n degrees of freedom, or in shorthand:

$$\frac{x_m/m}{x_n/n} = F_{m,n} \tag{9}$$

- The Fisher-F density with m and n degree of freedom is defined as

$$f_\theta(x) = \frac{\Gamma([m+n]/2)}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{m/2} \frac{x^{(m-2)/2}}{(1+(m/n)x)^{(m+n)/2}} \qquad x > 0 \tag{10}$$

where $\theta = [m, n]$ is a pair of positive integers.

- It should be noted that moments $E[X^k]$ of order greater than $k = n/2$ do not exist.

- A useful asymptotic relation for n large and $n >> m$ is $F_{m,n} \approx x_m$.

1.7 Cauchy Distribution

- The ratio of independent $N(0,1)$ r.v.'s U and V is called a standard Cauchy r.v. $X = U/V \sim C(0,1)$.

- It's density has the form

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2} \qquad x \in \mathbb{R} \tag{11}$$

- If $\theta = [\mu, \sigma]$ are location and scale parameters $(\sigma > 0) f_\theta(x) = f((x-\mu)/\sigma)$ is a translated and scaled version of the standard Cauchy density denoted $C(\mu, \sigma^2)$.

  Some properties of note:

- The Cauchy distribution has no moments of any (positive) integer order.

- The Cauchy distribution is the same as a Student-t distribution with 1 degrees of freedom.

1.8 Beta Distribution

- For $U \sim x_m$ and $V \sim x_n$ independent Chi-square r.v.s with m and n degrees of freedom, respectively, the ratio $X = U/(U+V)$ has a Beta distribution, or in shorthand

$$\frac{x_m}{x_m + x_n} = B(m/2, n/2) \tag{12}$$

  where $B(p, q)$ is a r.v. with Beta density having parameters $\theta = [p, q]$.

- The Beta density has the form

$$f_\theta(x) = \frac{1}{\beta_{r,t}} x^{r-1}(1-x)^{t-1} \qquad x \in [0,1] \tag{13}$$

  where $\theta = [r, t]$ and $r, t > 0$.

- Here $\beta(r, t)$ is the Beta function:

$$\beta_{r,t} = \int_0^1 x^{r-1}(1-x)^{t-1}dx = \frac{\Gamma(r)\Gamma(t)}{\Gamma(r+t)} \tag{14}$$

  Some useful properties:

- The special case of $m = n = 1$ gives rise to X an arcsin distributed r.v.

- $E_\theta[B(p,q)] = p/(p+q)$

- $\mathrm{var}_\theta(B(p,q)) = pq/((p+q+1)(p+q)^2)$

1.9 Reproducing Distributions

- A random variable X is said to have a reproducing distribution if the sum of two independent realizations, say X_1 and X_2, of X have the same distribution, possibly with different parameter values, as X.

- A Gaussian r.v. has a reproducing distribution:

$$N(\mu_1, \sigma_1^2) + N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \tag{15}$$

  which follows from the fact that the convolution of two Gaussian density functions is a Gaussian density function.

- Noting the stochastic representations of the Chi-square and non-central Chi-square distributions, respectively, it is obvious that they are reproducing distributions:

- $x_n + x_m = x_{m+n}$, if x_m, x_n are independent.

- $x_{m,\delta_1} + x_{n,\delta_2} = x_{m+n,\delta_1+\delta_2}$, if x_{m,δ_1} and $x_{n;\delta_2}$ are independent.

- The Chi square mixture, Fisher-F, and Student-t are not reproducing densities.

1.10 Fischer-Cochran Theorem

- This result gives a very useful tool for finding the distribution of quadratic forms of Gaussian random variables.

- *Theorem*: Let $X = [X1, \ldots, X_n]^T$ be a vector of iid. $N(0,1)$ rv's and let $\mathbf{A}$ be a symmetric idempotent matrix ($\mathbf{AA} = \mathbf{A}$) of rank p. Then

$$X^T \mathbf{A} X = x_p$$

2 Discrete Distributions

2.1 Binomial Distribution

- An experiment which follows a binomial distribution will satisfy the following requirements (think of repeatedly flipping a coin as you read these):
  - The experiment consists of n identical trials, where n is fixed in advance.
  - Each trial has two possible outcomes, S or F, which we denote "success" and "failure" and code as 1 and 0, respectively.
  - The trials are independent, so the outcome of one trial has no effect on the outcome of another.
  - The probability of success, $p = P(S)$, is constant from one trial to another. The random variable X of a binomial distribution counts the number of successes in n trials.
- The probability that X is a certain value x is given by the formula

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \tag{16}$$

where $0 \leq p \leq 1$, $x = 0, 1, \ldots, n$.

- Recall that the quantity, "n choose x" above is

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

- $E(x) = np$.
- $\text{var}(x) = np(1-p)$.

2.2 The Negative Binomial Distribution

- The negative binomial distribution is used when the number of successes is fixed and we are interested in the number of failures before reaching the fixed number of successes.
- An experiment which follows a negative binomial distribution will satisfy the following requirements:
  - The experiment consists of a sequence of independent trials.
  - Each trial has two possible outcomes, S or F.
  - The probability of success, $p = P(S)$, is constant from one trial to another.
  - The experiment continues until a total of r successes are observed, where r is fixed in advance.
  - A random variable X which follows a negative binomial distribution is denoted $X =$NB(r, p).
- Its probabilities are computed with the formula

$$P(X = x) = \begin{pmatrix} x + r - 1 \\ r - 1 \end{pmatrix} p^r (1 - p)^x \tag{17}$$

  where $0 \leq p \leq 1$, $x = 0, 1, 2 \ldots$.

- $E(x) = \frac{r(1-p)}{p}$.
- $\text{var}(x) = \frac{r(1-p)}{p^2}$.

2.3 Geometric Distribution

- The geometric distribution is a discrete distribution having probability function

$$P(X = x) = p(1 - p)^x \tag{18}$$

  for $0 \leq p \leq 1$, $x = 0, 1, 2 \ldots$.

- X is the number of failures before the first success in a sequence of independent Bernoulli trials.
- The geometric random variable X is the only discrete random variable with the memoryless property.
- $E(x) = \frac{1-p}{p}$.
- $\text{var}(x) = \frac{1-p}{p^2}$.

2.4 Poisson Distribution

- The Poisson distribution is most commonly used to model the number of random occurrences of some phenomenon in a specified unit of space or time.

  For example,

  - The number of phone calls received by a telephone operator in a 10-minute period.
  - The number of flaws in a bolt of fabric.
  - The number of typos per page made by a secretary.

- For a Poisson random variable, the probability that X is some value x is given by the formula

$$P(X = x) = \frac{\mu^x e^{-\mu}}{x!} \tag{19}$$

  where $x = 0, 1, 2 \ldots$, and μ is the average number of occurrences in the specified interval.

- $E(x) = \mu$.
- $\text{var}(x) = \mu$.