Statistical notes

Vahid Meghdadi

December 2008

1 Normal random variable

The probability density function is :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

The cdf function can be written using Q-function as:

$$F_X(x) = 1 - Q\left(\frac{m}{\sigma}\right)$$

The characteristic function of a random variable X is defined as:

$$\phi_X(u) = E\{e^{iuX}\}$$

For a normal distribution $X \sim \mathcal{N}(m, \sigma^2)$:

$$\phi_X(u) = E\{e^{iuX}\} = e^{ium - \frac{1}{2}\sigma^2 u^2}$$

Some properties of characteristic function:

- $\phi_X(0) = E\{e^{i0X}\} = 1$
- $|\phi_x(u)| \leq 1 = \phi_x(0)$
- $\phi_{aX+b} = e^{iub}\phi_X(0)$
- $\phi_X(-u) = \phi_X^*(u)$
- $E\{X^k\} = \frac{1}{i^k}\frac{d^k}{du^k}\phi_X(u)|_{u=0}$
- $-\frac{\partial}{\partial u} log\phi_X(u)|_{u=0} = iE\{X\}$
- $-\frac{\partial^2}{\partial u^2} log\phi_X(u)|_{u=0} = E\{X\}^2 - EX^2 = -VAR(X)$

2 Extension to random vector

Consider $\mathbf{X}$ a vector of size n containing n normal distributed random variables: $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{C})$. The pdf of X :

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n det(\mathbf{C})}} e^{-\frac{(\mathbf{x}-\mathbf{m})^T \mathbf{C}^{-1} (\mathbf{x}-\mathbf{m})}{2}}$$

where covariance matrix $\mathbf{C}$ is:

$$\mathbf{C} = E\{[X - E\{X\}][X - E\{X\}]^T\} = \mathbf{C}^T$$

Note that $(2\pi)^n det(\mathbf{C}) = det(2\pi\mathbf{C})$

We can write the characteristic function of this multivariate normal vector:

$$\phi_{\mathbf{X}}(\mathbf{u}) = E(e^{j\mathbf{u}^T\mathbf{X}}) = e^{i\mathbf{m}_X^T\mathbf{u} - \frac{1}{2}\mathbf{u}^T\mathbf{C}_X\mathbf{u}}$$

Here there is no need to inverse the covariance matrix.

We can write:

$$EX_j = i^{-1}\frac{\partial}{\partial u_j} log\phi_{\mathbf{X}}(\mathbf{Y})|_{\mathbf{u}=0} = m_j$$

$$EX_iX_j = i^{-2}\frac{\partial^2}{\partial u_j u_k} log\phi_{\mathbf{X}}(\mathbf{Y})|_{\mathbf{u}=0} = m_j m_k + C_{jk}$$

If X_i are independent :

$$\phi_{\mathbf{X}}(\mathbf{u}) = \prod_{i=1}^{n} \phi_{X_i}(u_i)$$

3 Linear transformation

Consider the transformation $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ where $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{R}_X)$

$\mathbf{Y}$ is a multivariate normal vector with:

$$E(\mathbf{Y}) = \mathbf{Am} + \mathbf{b}$$

$$\mathbf{R}_Y = \mathbf{AR_XA}^T + \mathbf{bb}^T + \mathbf{bm}^T\mathbf{A}^T + \mathbf{Amb}^T$$

and $\mathbf{C}_Y = \mathbf{AC}_X + \mathbf{A}^T$.
We can also writ the characteristic function of $\mathbf{Y}$:

$$\phi_{\mathbf{Y}}(\mathbf{u}) = Ee^{i\mathbf{u}^T\mathbf{Y}} = e^{i\mathbf{b}^T\mathbf{u}}\phi_X(\mathbf{A}^T\mathbf{u})$$

Generalized transformation of a random vector
$\mathbf{X}$ is an $\mathcal{R}^n$-valued random vector and $G(\mathbf{x})$ is an invertible vector valued function of $\mathbf{x} \in \mathcal{R}^n$. So $\mathbf{Y} = G(\mathbf{x})$ is a random vector. Knowing that the probability

density function of $\mathbf{X}$ is $f_{\mathbf{X}}(\mathbf{x})$ the question is : what is the *pdf* of $f_{\mathbf{Y}}(\mathbf{y})$?

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} g_1(x_1, \cdots x_n) \\ \vdots \\ g_n(x_1, \cdots x_n) \end{bmatrix} \tag{1}$$

Because G^{-1} exists, $H(\mathbf{y} \doteq G^{-1}(\mathbf{y})$, so

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} h_1(y_1, \cdots y_n) \\ \vdots \\ h_n(y_1, \cdots y_n) \end{bmatrix} \tag{2}$$

Hypothesis H is continuous with continuous partial derivatives.

$$dH(\mathbf{y}) \doteq \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_n}{\partial y_1} & \cdots & \frac{\partial h_n}{\partial y_n} \end{bmatrix} \tag{3}$$

To compute *pdf* of $\mathbf{Y}$, we begin with $P(\mathbf{Y} \in \mathcal{C}) = P(G(\mathbf{x}) \in \mathcal{C})$. It is convenient to define the set $\mathcal{B} \doteq \{\mathbf{x} : G(\mathbf{x}) \in \mathcal{C}\}$

4 Gamma function

The gamma function is defined as

$$\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt$$

for $p > 0$. There are some useful properties:

$$\Gamma(p) = (p-1)! \qquad p \texttt{ an integer and } p > 0$$
$$\Gamma(1/2) = \sqrt{\pi}, \qquad \Gamma(3/2) = \frac{\sqrt{\pi}}{2}$$

$$\Gamma(n + 1/2) = \sqrt{\pi} \frac{\prod_{k=1}^{n} (2k-1)}{2^n}$$

5 Chi-Square Distribution

Let X be a zero-mean Gaussian-distributed random variable. The variable $Y = X^2$ has a chi-squared distribution with:

$$p_Y(y) = \frac{1}{\sqrt{2\pi y \sigma}} e^{-y/2\sigma^2} u(y)$$

The CDF cannot be expressed in closed form. How ever the characteristic function has a closed form as:

$$\psi(jv) = \frac{1}{(1 - j2v\sigma^2)^{1/2}}$$

Now suppose the RV Y defined as:

$$Y = \sum_{i=1}^{n} X_i^2$$

The characteristic function of Y is:

$$\psi_Y(jv) = \frac{1}{(1 - j2v\sigma^2)^{n/2}}$$

The inverse transform of this characteristic function yields the pdf:

$$p_Y(y) = \frac{1}{\sigma^n 2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2\sigma^2} u(y)$$

which is called the chi-square distribution with n degrees of freedom. The CDF can be calculated analytically:

$$F_Y(y) = 1 - e^{-y/2\sigma^2} \sum_{k=0}^{m-1} \frac{1}{k!} \left(\frac{y}{2\sigma^2}\right)^k$$

where $m = n/2$. There are also the moments of chi-square distribution as follows:

$$\mathrm{E}(Y) = n\sigma^2$$
$$\mathrm{E}(Y^2) = 2n\sigma^4 + n^2\sigma^4$$
$$\sigma_y^2 = 2n\sigma^4$$

If the Gaussian random variables are not zero-mean, the pdf of resulting distribution $P_Y(y)$ can be calculated in terms of Marcum's Q-function (see Proakis).

6 Rayleigh distribution

Suppose there are two i.i.d. RV each with normal distribution: $X_1, X_2 \sim \mathcal{N}(0, \sigma^2)$. Let's define $Y = (X_1 + X_2)^{1/2}$. The RV Y follows the Rayleigh distribution as follows:

$$f_X(x) = \frac{x}{\sigma^2} exp\left(\frac{-x^2}{2\sigma^2}\right) U(x)$$

The mean and variance are:

$$m_x = \sigma(\pi/2)^{1/2}$$
$$\sigma_x^2 = \sigma^2(2 - \pi/2)$$

The CDF is:

$$F_X(x) = 1 - e^{-x^2/2\sigma^2} U(x)$$

7 Nakagami-m distribution

The pdf of Nakagami-m distribution is defined by:

$$p_\alpha(\alpha) = \frac{2m^m \alpha^{2m-1}}{\Omega^m \Gamma(m)} exp\left(-\frac{m\alpha^2}{\Omega}\right) U(\alpha) \quad (4)$$

where m is its parameter which ranges from $1/2$ to ∞ and $\Omega = \bar{\alpha^2}$. For $m = 1/2$ this distribution reduces to one sided Gaussian ($\Gamma(1/2) = \sqrt{\pi}$) and for $m = 1$ it is the Rayleigh distribution ($\Gamma(1) = 1$).

Having the attenuation parameter as Nakagami-m, the signal to noise ratio at the receiver follows gamma distribution:

$$p_\gamma(\gamma) = \frac{m^m \gamma^{m-1}}{\bar{\gamma}^m \Gamma(m)} exp\left(-\frac{m\gamma}{\bar{\gamma}}\right) U(\gamma) \quad (5)$$

This is obtained using the following formulas (ref: Digital Communication over Fading Channel, by Simon)

$$p_\gamma(\gamma) = \frac{p_\alpha\left(\sqrt{\frac{\Omega\gamma}{\bar{\gamma}}}\right)}{2\sqrt{\frac{\gamma\bar{\gamma}}{\Omega}}}$$

where $\gamma = \alpha^2 E_s/N_0$.

One can show that the MGF of γ is given by:

$$M_\gamma(s) = \left(1 - \frac{s\bar{\gamma}}{m}\right)^{-m}$$

and therefore:

$$\mathrm{E}[\gamma^k] = \frac{\Gamma(m+k)}{\Gamma(m)m^k}\bar{\gamma}^k$$

Note that the Nakagami-m distribution often gives the best fit to land-mobile and indoor-mobile multipath propagation.

8 Bounds

8.1 Chernoff Bound

The problem is where we have a sum of *i.i.d.* random variables. It is difficult to calculate the distribution of the sum because of the number of variable can be important. Therefore upper bound calculations can be used. To be continued.,