Point Estimation

Error-Control Coding and the Binomial Distribution

Hans Georg Schaathun

Session 28 January 2015
Updated 20th January 2015

1 Error rates

We are interested in the probability that a word is decoded incorrectly, when it is transmitted over a channel with error control coding. You tested this in the last session with the BSC and the Hamming code. Conducting m tests and counting the number of decoding errors X, we can calculate the *error rate* X/m.

Exercise 1 *Create an m-file which runs m tests where a random word is coded with the Hamming code, transmitted over BSC(p), and decoded. The input is the numbers m and p, and the output is the error rate X/m where X is the number of wrongly decoded words.*

Exercise 2 *Run the function from the last exercise for $p = 0.1$ and $m = 100$. What is the error rate? What can you say about the error probability?*

2 The standard deviation

The m-file from Exercise 1 is a random function. It will not (normally) give you the same answer twice. Let's investigate the variation.

Exercise 3 *Run the function from Exercise 1 100 times with the same parameters as in Exercise 2, to get a vector of 100 observations of the error rate. Have a brief look at the results. What is your gut feeling about the error probability?*

Exercise 4 *The last exercise gave you a sample of 100 observations of the error rate. Calculate the sample mean and the sample standard deviation. What do these measures tell you about the error probability?*

For for any stochastic variable X and scalar α, we have

$$E(\alpha X) = \alpha E(X), \tag{1}$$
$$\mathsf{var}(\alpha X) = \alpha^2 \,\mathsf{var}(X). \tag{2}$$

For any ensemble of independently distributed stochastic variables X_i, we have

$$E\big(\sum X_i\big) = \sum E(X_i), \tag{3}$$
$$\mathsf{var}\,\big(\sum X_i\big) = \sum \mathsf{var}(X_i). \tag{4}$$

Table 1: Some formulæ.

3 Standard error

The number X of word errors when n words are transmitted on a BSC is binomially distributed, $X \sim B(n,p)$. The parameter p is called the (word) error probability or the probability of decoding error. For most codes, it is not feasible to calculate p analytically.

The error rate X/n is an *estimator* for the error probabilty, and we denote it by $\hat{p} = X/n$. The estimator is a stochastic variable, and an observation of $\hat{p}$ is called an *estimate.*

What do we know about the probability distribution on $\hat{p}$*?*

Exercise 5 *Find expressions for the following parameters as functions of* p*:*

1. *the expected value* $E(\hat{p})$
2. *the variance* $\mathsf{var}(\hat{p})$
3. *the standard deviation* $\mathsf{StdDev}(\hat{p})$ *(i.e. the standard error).*

4 Estimating the Standard Error

Since p is unknown, we cannot use Exercise 5 to calculate the standard error. In Exercise 4 you calculated an estimate for the standard error.

Exercise 6 *What is the estimate for the standard error calculated in Exercise 4?*

In Exercise 4 you calculated $\hat{p}$ 100 times to estimate the standard error. There is a better way to estimate the standard error, based on estimating p only once. We can use Exercise 5, replacing the parameter p with its estimate.

Exercise 7 *Run your function from Exercise 1 once to generate an estimate $\hat{p}$ for the error probability.*

1. *What is the estimated error probability?*

2. *What estimate do you get for the standard error?*

Exercise 8 *Compare the standard error estimates from Exercises 4 and 7. Are the results as expected?*

Exercise 9 *Imagine that you consider using the Hamming code and the BSC in a practical system, and need some assurance about the error probability. Based on your answers to Exercise 7, what do you feel you can say about the error probability?*