SHANNON'S THEOREM AND THE

BINOMIAL RANDOM VARIABLE

Martin Erickson

Northeast Missouri State University

Claude E. Shannon's 1948 theorem on coding over a noisy channel states that information can be transmitted with arbitrarily high accuracy at any rate below the channel capacity. Channel and capacity are defined in the references below. The proof requires a combinatorial or probabilistic lemma about binomial coefficients $C(n, j)$.

Lemma. If $0 < x < \frac{1}{2}$ and

$$H(x) = -x \log x - (1-x) \log(1-x),$$

then

$$\sum_{j=0}^{[nx]} C(n, j) < 2^{nH(x)}.$$

The logarithms are base two and the quantity $H(x)$ is the entropy of the channel, measured in bits. All previously published proofs of the lemma use subtle approximation methods (often Stirling's approximation to the factorial function). The following new proof presumes only the binomial theorem.

Proof. With $x + y = 1$ the binomial theorem and two simple inequalities yield

$$1 = \sum_{j=0}^{n} C(n, j) x^j y^{n-j} > \sum_{j=0}^{[nx]} C(n, j) x^j y^{n-j}$$

$$> \sum_{j=0}^{[nx]} C(n, j) x^j y^{n-j} (x/y)^{nx-j} = \sum_{j=0}^{[nx]} C(n, j) x^{nx} y^{ny}.$$