8.1. **a.** Prove that if a real-valued random variable has finite variance, then it also has finite entropy.

b. If a real-valued random variable has finite entropy, does this imply that it has finite variance?

8.2. Prove that, of all probability density functions defined on the interval $[0, 1]$, the uniform density function has the largest differential entropy.

8.3. Suppose that a gaussian-noise channel with average power constraint has the noise power density spectrum

$$N(f) = \frac{N_0}{2}$$

and an input filter

$$|H(f)|^2 = \frac{1}{1 + (f/f_0)^2}.$$

Find $C(S)$.

8.4. The *Pareto probability density function* is given by

$$p(x) = \frac{ak^a}{x^{a+1}} \quad \text{for} \quad x \geq k,$$

and $p(x) = 0$ for $x < k$ where the parameters k and a are positive real numbers. Prove that the differential entropy of a Pareto random variable is

$$H(p) = \log_e \frac{k}{a} + 1 + \frac{1}{a}.$$

Prove that the variance of a random variable with Pareto probability density function is

$$\sigma^2 = \frac{k^2 a}{(a-1)^2(a-2)}$$

for $a > 2$.

8.5. The bivariate gaussian probability density function is given by

$$p(x, y) = \frac{1}{2\pi|\boldsymbol{\Sigma}|} e^{-(x,y)\boldsymbol{\Sigma}^{-1}(x,y)^T}$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}^{-1}$$

is the covariance matrix of the bivariate random variable. Give expression for the marginals $p(x)$ and $p(y)$, and the conditionals $p(x|y)$ and $p(y|x)$ in terms of a, b, and c.
Hint: $ax^2 + 2bxy + cy^2 = (a - b^2/c)x^2 + (cy + bx)^2/c$.

8.6. The capacity of an additive gaussian-noise waveform channel is given by the water-filling formulas

$$C(S) = \frac{1}{2}\int_{-\infty}^{\infty} \max\left[0, \log \frac{\theta}{N(f)/|H(f)|^2}\right] df$$
$$S = \int_{-\infty}^{\infty} \max\left[0, \theta - \frac{N(f)}{|H(f)|^2}\right] df,$$

where $H(f) = 0$ for $|f| \geq W$. An adversary, called a "jammer," intends to interfere with the communication system by creating additive gaussian noise, independent of the signal, having spectrum $N(f)$.

a. Could the jammer use a different, nongaussian noise (in which case the given formula does not apply), to obtain a smaller value of $C(S)$? Why?

b. How should a gaussian jammer choose $N(f)$ subject to the power constraint

$$\int_{-\infty}^{\infty} N(f) df \leq J,$$

so as to minimize the capacity?

c. If the jammer can use noise that is dependent on the signal, can $C(S)$ be made smaller? Give a reason.

8.7. Two independent additive gaussian-noise subchannels are in parallel. They have noise variance $N_1 = 1$ unit and $N_2 = 2$ units, respectively. A total transmitted power S of three units is to be distributed between the two channels.

a. How should the power be distributed to maximize total capacity? What is that total capacity?

b. How should the power be distributed so that the two channels have the same individual capacity? What is the total capacity?

8.8. A single-input, two-output channel (for example, a channel with dual receiving antennas) but a single demodulator has an input X and an output (Y_1, Y_2), given by

$$Y_1 = X + Z_1$$
$$Y_2 = X + Z_2,$$

where Z_1 and Z_2 are memoryless, zero-mean gaussian random variables with covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}.$$

The input power is limited to S.

a. Find the capacity if $\rho = 1$.

b. Find the capacity if $\rho = 0$.

c. Find the capacity if $\rho = -1$.

d. Find the capacity for arbitrary ρ.