MASSACHUSETTS INSTITUTE OF TECHNOLOGY

2.717J/MAS.857J Optical Engineering **Spring '02**

Problem Set #2 Posted Feb. 13, 2002 — Due Wednesday Feb. 20, 2002

1. **How to emulate a perfect coin.** Given a biased coin such that the probability of heads (H) is α, we emulate a perfect coin as follows: Throw the biased coin twice; interpret HT (T =tails) as success and TH as failure; if neither event occurs repeat the throws until a decision is reached.

   1.a) Show that this model leads to Bernoulli trials with $p = 1/2$.

   1.b) Find the distribution and the expectation value of the number of throws required to reach a decision.

2. **Birthdays.** For a group of n people find the expected number of days of the year which are birthdays of exactly k people. Assume the year is 365 days long and that all the arrangements are equally probable. What is the result for $n = 23$ (the number of players in two opposing soccer teams plus the referee) and $k = 2$? Do you find that surprising?

3. **Misprints.** A book of n pages contains on average λ misprints per page. Estimate the probability that at least one page will contain more than k misprints.

4. **Detection threshold.** We seek to determine if a tumor is present in tissue from the voltage U measured between two strategically placed electrodes. In the absence of tumor, U is Gaussian with mean V_1 and variance σ^2; *i.e.*, the "prior" distribution is

$$p_U(u \mid \text{no tumor}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(u - V_1)^2}{2\sigma^2}\right\}.$$

   In the presence of a tumor, U is Gaussian with mean $V_2 > V_1$ and same variance σ^2; *i.e.*

$$p_U(u \mid \text{tumor}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(u - V_2)^2}{2\sigma^2}\right\}.$$

   We seek a "detection threshold" V_0 such that if $U > V_0$ we conclude that a tumor is present; whereas if $U < V_0$ we conclude that there is no tumor. Clearly, our decision is in error if **(i)** we concluded that there is no tumor whereas in actuality a tumor is present, *i.e.* a "miss;" **(ii)** we concluded that there is a tumor whereas in actuality there is no tumor present, *i.e.* a "false alarm." We define the probability of error (PE) as the sum of the probability of a miss and the probability of a false alarm.

4.a) Show that the PE is minimized if we select

$$V_0 = \frac{V_1 + V_2}{2}.$$

4.b) Using the optimum threshold, calculate the PE in terms of the "error function"

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \mathrm{e}^{-t^2} \mathrm{d}t.$$

Notes: (1) The above-described process of selecting a detection threshold is known as "Bayes decision." (2) The erf definition above is after Abramowitz & Stegun, *Handbook of Mathematical Functions*, Dover 1972 (p. 297). The constants and integral limits are sometimes defined differently in the literature.

5. Normalization. Let $\{X_k\}$ be a sequence of mutually independent random variables with a common distribution. Suppose that the X_k assume only positive values and that $\mathrm{EV}\{X_k\} = \overline{x}_k = a$ and $\mathrm{EV}\left\{X_k^{-1}\right\} = b$ exist. Let

$$S_n = X_1 + \ldots + X_n.$$

Prove that $\mathrm{EV}\{S_n^{-1}\}$ is finite and that

$$\mathrm{EV}\left\{\frac{X_k}{S_n}\right\} = \frac{1}{n} \qquad \text{for } k = 1, \ldots, n.$$

6. Unbiased estimator. Let $X_1, \ldots, X_n$ be mutually independent random variables with a common distribution; let its mean be μ, its variance σ^2. Let

$$\overline{X} = \frac{X_1 + \ldots + X_n}{n}.$$

Prove that

$$\sigma^2 = \frac{1}{n-1} \mathrm{EV}\left\{\sum_{k=1}^{n} \left(X_k - \overline{X}\right)^2\right\}.$$

(Note: In statistics, $\overline{X}$ is called an *unbiased estimator* of $\overline{x} = \mathrm{EV}\{X\}$, and $\sum \left(X_k - \overline{X}\right)^2 / (n-1)$ is an *unbiased estimator* of σ^2.