University of Illinois **Fall 2011**

ECE 313: Problem Set 4: Solutions
Confidence intervals, ML parameter estimation, Bernoulli processes, Poisson Distribution

1. **[Confidence Interval]**

   (a) The error count E can be obtained by running n independent trials of a Bernoulli random variable with parameter p_e. Thus, E is a binomial random variable with parameters (n, p_e).

   (b) The probability of error p_e should lie in the range $\hat{p}_e - 10^{-5} \leq p_e \leq \hat{p}_e + 10^{-5}$ in order to meet a $\pm 10\%$ tolerance. For a 99% confidence level, i.e., $a = 10$,

   $$\frac{a}{2\sqrt{n}} = 10^{-5}$$

   Thus, $n = 25 \times 10^{10}$ bits are needed.

   (c) The desk-top CPU executes 4×10^9 cycles/sec. Thus, $\frac{4\times 10^9}{40} = 10^8$ bits are simulated per sec. As there are $n = 25 \times 10^{10}$ bits, one simulation run will take $\frac{25\times 10^{10}}{60\times 10^8} = 41.6$ minutes.

   (d) The simulation run takes 600 seconds, which corresponds to 24×10^{11} clock cycles. Thus, $n = \frac{24\times 10^{11}}{40} = 6 \times 10^{10}$ bits can be simulated, leading to

   $$\frac{a}{2\sqrt{6 \times 10^{10}}} = 10^{-5} \Rightarrow a = 2\sqrt{6}$$

   and thus a confidence level of

   $$1 - \frac{1}{a^2} = 1 - \frac{1}{24} = 0.958$$

   i.e., a 95.8% confidence level. The designer considers this loss in confidence in the initial runs as a reasonable compromise in order to get the project completed on time.

2. **[Geometric Random Variables]**

   (a) $p_{\mathbf{Y}}(k) = p(1-p)^{k-1}$ for $k \geq 1$.

   (b) $\mathrm{E}[\mathbf{Y}] = \frac{1}{p} = \begin{cases} 1.0526 & \text{A student} \\ 6.6667 & \text{C student} \end{cases}.$

   (c) $\mathrm{E}[\mathbf{Y}] = \frac{1}{p} < 5$, so $p > 0.2 > 0.15$. Thus the C student might risk not getting a job!

   (d) P(an A student does not get an invitation in 5 trials)$=\sum_{k=6}^{\infty} p(1-p)^{k-1} = (1-p)^5 \sum_{k'=0}^{\infty} p(1-p)^{k'} = (1-p)^5 = (1-0.95)^5 = 3.125 \times 10^{-7}$.
   It can also be observed directly that P(an A student does not get an invitation in 5 trials) $= (1-p)^5$ by the independence of each trial.
   P(C gets an invitation in 5 trials)$=1-(1-p)^5 = 1-(1-0.15)^5 = 0.5563$.

3. **[Maximum-likelihood Estimation]**

   (a) Let's say $\hat{n}_{\mathrm{ML}} < 10$. Then the set of values that $\mathbb{X}$ can take on does not include 10 or any values higher than that, so it is impossible to obtain 10 as the value of $\mathbb{X}$, as it is not in the alphabet $\{1, 2, ..., \hat{n}_{\mathrm{ML}}\}$.

   (b) Given $\mathbb{X} = 10$, the likelihood function is $L(n) = p_{\mathbb{X}}(10) = \frac{20}{n(n+1)}$ which is a *decreasing* function of n. The maximum thus occurs at the minimum value for n. Thus, $\hat{n}_{\mathrm{ML}} = 10$.

(c) i. Compute the value of $p_{\mathbb{X}}(10)$ for $n = 10, 11, 12, \ldots$ and find the maximum-likelihood estimate $\hat{n}_{\text{ML}}$ numerically.

It is easy to compute the values of the likelihood function $L(n) = p_{\mathbb{X}}(10) = \frac{2(n+1-10)}{n(n+1)}$ for $n \geq 10$. The maximum value is $0.05263158\ldots = \frac{1}{19}$ for $n = 18$ and also for $n = 19$. Hence, we conclude $\hat{n}_{\text{ML}}$ is one of $\{18, 19\}$.

ii. Now suppose that $\mathbb{X}$ has value i. Find $\hat{n}_{\text{ML}}$ as a function of i and verify that when $i = 10$, your function gives the same value for $\hat{n}_{\text{ML}}$ as you found in part (c)(i).

We look at the ratio $\dfrac{L(n)}{L(n-1)} = \dfrac{\frac{2(n+1-i)}{n(n+1)}}{\frac{2(n-i)}{(n-1)n}} = \dfrac{n^2 - ni + i - 1}{n^2 - ni + n - i}$ which is greater than 1, equal to 1, and less than 1, according as $i - 1$ is greater than, equal to, or less than $n - i$, that is, according as n is less than, equal to, or greater than $2i - 1$. In other words,

$$L(i) < L(i+1) < L(i+2) < \cdots < L(2i-2) = L(2i-1) > L(2i) > L(2i+1) > \cdots$$

so that if we observe that $\mathbb{X} = i$, then the maximum-likelihood estimate of n is $\hat{n}_{\text{ML}} = 2i - 2$ or $\hat{n}_{\text{ML}} = 2i - 1$. For $i = 10$, this result matches the one we obtained by computation.

Notice the difference in the maximum-likelihood estimates of n for the increasing-ramp and decreasing-ramp pmfs. If we observe that $\mathbb{X} = i$, we estimate $\hat{n}_{\text{ML}} = i$ for increasing-ramp pmfs, while for decreasing-ramp pmfs, the maximum-likelihood estimate is $\hat{n}_{\text{ML}} = 2i-2$ or $\hat{n}_{\text{ML}} = 2i-1$, nearly twice the observed value of $\mathbb{X}$.

4. **[Poisson and Binomial Distributions]**

(a) On average, $\mathsf{E}[\mathbb{X}] = 105 \times 0.9 = 94.5$ passengers show up for the flight.

(b) $P\{\mathbb{X} \leq 100\} = 1 - P\{\mathbb{X} > 100\}$
$= 1 - P\{\mathbb{X} = 101\} - P\{\mathbb{X} = 102\} - P\{\mathbb{X} = 103\} - P\{\mathbb{X} = 104\} - P\{\mathbb{X} = 105\}$
$= 1 - \binom{105}{101}(0.9)^{101}(0.1)^4 - \binom{105}{102}(0.9)^{102}(0.1)^3 - \binom{105}{103}(0.9)^{103}(0.1)^2 - \binom{105}{104}(0.9)^{104}(0.1)^1 - \binom{105}{105}(0.9)^{105}(0.1)$
$= 1 - \binom{105}{4}(0.9)^{101}(0.1)^4 - \binom{105}{3}(0.9)^{102}(0.1)^3 - -\binom{105}{2}(0.9)^{103}(0.1)^2 - \binom{105}{1}(0.9)^{104}(0.1)^1 - \binom{105}{0}(0.9)^{105}(0.1)$
$= 0.9832\ldots$

(c) If $\mathbb{X}$ is a binomial random variable with parameters (n, p), then $\mathbb{Y} = n - \mathbb{X}$ is a binomial random variable with parameters $(n, 1 - p)$.

(d) $P\{\mathbb{Y} \geq 5\} = 1 - P\{\mathbb{Y} = 0\} - P\{\mathbb{Y} = 1\} - P\{\mathbb{Y} = 2\} - P\{\mathbb{Y} = 3\} - P\{\mathbb{Y} = 4\}$
$= 1 - \exp(-10.5)\left[1 + \frac{10.5}{1!} + \frac{(10.5)^2}{2!} + \frac{(10.5)^3}{3!} + \frac{(10.5)^4}{4!}\right] = 0.9789\ldots$ which is close enough to $0.9832\ldots$ for gummint work.

5. **[Inferring true performance on multiple-choice examinations]**

(a) We are modelling the guesses as independent trials, and we are not allowing for other possibilities such as on some questions, the student can eliminate one or more alternatives and thus improve chances of getting the right answer to $\frac{1}{4}$ or $\frac{1}{3}$ etc.

(b) The likelihood that we observe $\mathbb{W} = n$ is $\binom{N-K}{n}(0.8)^n(0.2)^{N-K-n}$ for $n = 0, 1, \ldots, N - K$.

(c) For given N and n, the likelihood of part (b) is a function, say $L(K)$ of K. We have that

$$\frac{L(K)}{L(K-1)} = \frac{\binom{N-K}{n}(0.8)^n(0.2)^{N-K-n}}{\binom{N-(K-1)}{n}(0.8)^n(0.2)^{N-(K-1)-n}} = \frac{N - (K-1) - n}{0.2(N - (K-1))} \geq 1 \text{ iff } K \leq N - 1.25n + 1.$$

Thus, the likelihood is maximum when K has value $\hat{K}_{\text{ML}} = \lfloor N - 1.25n + 1 \rfloor$.

(d) If n is *not* a multiple of 4, then the analysis of part (b) shows that $L(\hat{K}_{\text{ML}})$ is the unique maximum of the likelihood, and the examiner's estimate $\hat{K}_{\text{GP}} = N - n - \lfloor 0.25n \rfloor$ equals the maximum likelihood estimate. When n *is* a multiple of 4, then $L(\hat{K}_{\text{ML}}) = L(\hat{K}_{\text{ML}} - 1)$, and thus both $\hat{K}_{\text{ML}}$ and $\hat{K}_{\text{ML}} - 1$ are legitimate maximum-likelihood estimates of the unknown quantity. Note that in this case, the examiner's estimate $\hat{K}_{\text{GP}} = \hat{K}_{\text{ML}} - 1$ and thus the examiner *is* using a maximum-likelihood estimate (she is just a tough grader!).

For $N = 100$ and $K = 90$, $\mathbb{W}$ can take on values $0, 1, \ldots, 10$. The value of $\mathbb{W}$ most likely to occur is 8. In this case, $\hat{K}_{\text{ML}} = 91$ while $\hat{K}_{\text{GP}} = 90$ and so the examiner *does* estimate K correctly. On the other hand, if $\mathbb{W} = 4$, the examiner estimates K to be $\hat{K}_{\text{GP}} = 95$ (erring on the side of caution), and if $\mathbb{W} = 10$, meaning that the 90 questions that the student knew the answers to were correctly, but *all* the remaining 10 questions were answered incorrectly, then $\hat{K}_{\text{GP}} = 88$, ouch!