STAT 598G Spring 2011
Jeremiah Rounds and Sergey Kirshner
Homework 2 (Written Exercises)

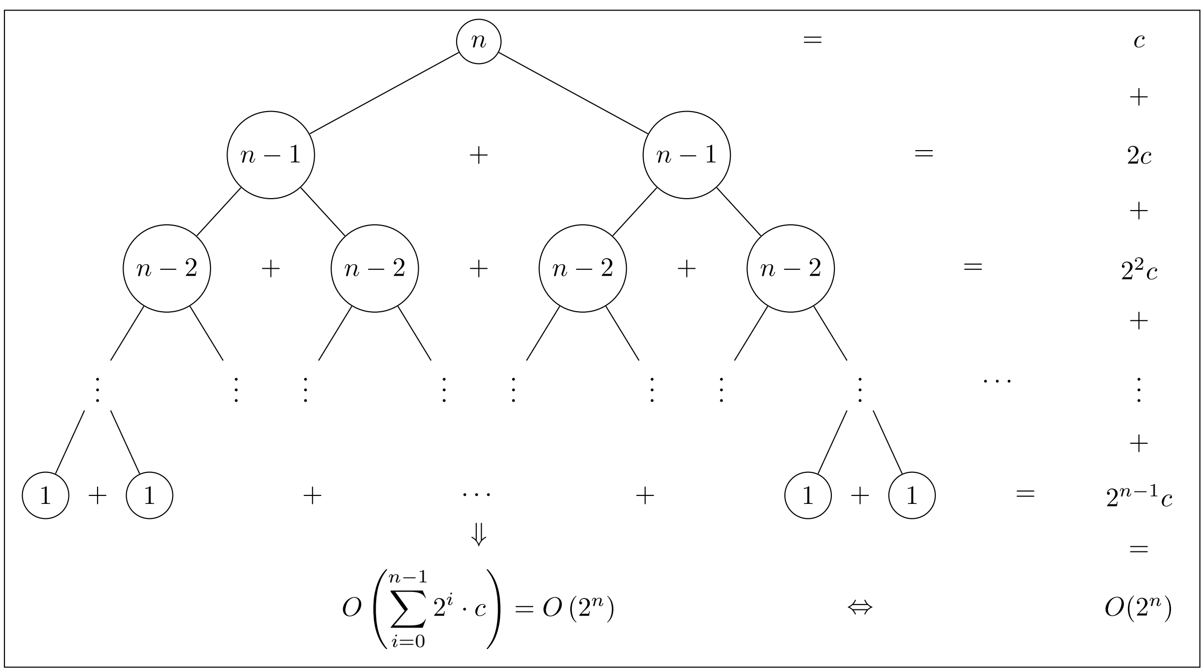
February 14, 2011

1. (2 pts) `http://learning.stat.purdue.edu/wiki/_media/courses/sp2011/598g/intro.tar.gz` contains file `recursion.c` which did not examine in class. What is the recurrence that is being solved by this program? What is the order of the running time complexity for the function `f(n)`?

   Solution: The program is computing the recurrence $f(n) = f(n-1) + f(n-2)$ for $n \in \mathbb{N}$. The base case for the recurrence is $f(1) = 1$ and $f(0) = 0$. The second part is tricky. Denote by $T(n)$ the number of operations needed to compute $f(n)$. Checking for $n > 1$ and adding two numbers takes constant time, say total of c. (We can assume that returning n if $n \leq 1$ takes the same time c.) Then $T(n) = T(n-1) + T(n-2) + c$, the same type of recurrence f is solving!

   Instead of finding the exact order Θ, we will instead find the asymptotic upper bound O. Note that $T(n) \geq T(n-1)$ for $n \geq 1$. Therefore

$$T(n) = T(n-1) + T(n-2) + c \leq 2T(n-1) + c.$$



   The illustration in the figure above suggests that $T(n) \in O(2^n)$. (Enough for full credit.) 2^n is however a loose asymptotic upper bound. One can actually show that $T(n) \in \Theta\left(\left(\frac{\sqrt{5}+1}{2}\right)^n\right)$ (bonus part of programming homework 2, exercise 1).

2. (1 pts) Exercise 2.1 (Robert and Casella, 2009): For an arbitrary random variable X with cdf F, define the generalized inverse of F by

$$F^{-}(u) = \inf\{x : F(x) \geq u\}.$$

Show that if $U \sim \mathsf{Unif}(0,1)$, then $F^{-}(U)$ is distributed like X.

Solution: Let $Y = F^{-}(U)$. Then

$$P(Y \leq y) = P\left(F^{-}(U) \leq y\right).$$

Assume $u \in (0,1)$. For any (y,u), $F^{-}(u) \leq y$ implies $F(y) \geq u$ (from the definition of $F^{-}(u)$). On the other hand, for any (y,u) such that $F(y) \geq u$, $F^{-}(u) \leq y$ (again from the definition of $F^{-}(u)$). Thus $\{(y,u) : F^{-}(u) \leq y\} = \{(y,u) : F(y) \geq u\}$. So

$$P(Y \leq y) = P\left(F^{-}(U) \leq y\right) = P(F(y) \geq U) = P(U \leq F(y)) = F(y) = P(X \leq y).$$

3. (1 pts) Gumbel distribution has the cumulative distribution function

$$F(x;\mu,\beta) = e^{-e^{-(x-\mu)/\beta}}, \ x \in \mathbb{R}.$$

Assuming one can draw samples from $\mathcal{U}(0,1)$, describe (and prove) how to draw samples from F.

Solution: We apply the inverse transformation method to sample from F. To do so, we note that F is continuous and is therefore invertible on its range of $(0,1)$. Now, assume $U \sim \mathsf{Unif}(0,1)$. Now, we compute the inverse transformation

$$u = F(x) = e^{-e^{-(x-\mu)/\beta}} \Leftrightarrow -\ln u = e^{-(x-\mu)/\beta} \Leftrightarrow -\ln(-\ln u) = \frac{x-\mu}{\beta} \Leftrightarrow x = \mu - \beta \ln(-\ln u).$$

Thus $F^{-1}(u) = \mu - \beta \ln(-\ln u)$ for $u \in (0,1)$. One can obtain samples $x \sim F$ by drawing $u \sim \mathsf{Unif}(0,1)$, and then setting $x = \mu - \beta \ln(-\ln u)$.

4. (1 pts) Prove that the Box-Muller procedure produces two independent standard normal random variables.

Solution: We pick up where the notes left off. Assuming $u_1, u_2 \overset{iid}{\sim} \mathsf{Unif}(0,1)$, $f_{U_1U_2}(u_1,u_2) = 1$ if $u \in (0,1)$, $v \in (0,1)$ and 0 otherwise,

$$f_{Z_1Z_2}(z_1,z_2) = f_{U_1U_2}(h_1(z_1,z_2), h_2(z_1,z_2)) \left|\det J(z_1,z_2)\right| = \frac{1}{2\pi}\exp\left(-\frac{1}{2}\left(z_1^2 + z_2^2\right)\right)$$

for z_1, z_2 s.t. $h_1(z_1,z_2) \in (0,1)$, $h_2(z_1,z_2) \in (0,1)$ which it is for $(z_1,z_2) \in \mathbb{R}^2$. Computing marginals of $f_{Z_1Z_2}(z_1,z_2)$ yields $Z_1, Z_2 \sim \mathcal{N}(0,1)$. Finally, we need to notice that $f_{Z_1}(z_1) f_{Z_2}(z_2) = f_{Z_1Z_2}(z_1,z_2)$ for all $(z_1,z_2) \in \mathbb{R}^2$, so Z_1 and Z_2 are independent.

5. (2 pts) Describe an Accept-Reject sampling algorithm for Rayleigh random variables, and find its probability of acceptance. Rayleigh's pdf is

$$f(x;\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \ x \geq 0$$

with a parameter $\sigma > 0$.

Solution: There are several possible solution. Here, we will use $g(x) = e^{-x},\ x \geq 0$, the density for the standard normal random variable as the proposal distribution for Accept-Reject sampler. Then $r(x) = f(x)/g(x) = \frac{x}{\sigma^2} e^{x - \frac{x^2}{2\sigma^2}},\ x \geq 0$.

$$h'(x) = \frac{1}{\sigma^2} e^{x - \frac{x^2}{2\sigma^2}} + \frac{x}{\sigma^2} e^{x - \frac{x^2}{2\sigma^2}} \left(1 - \frac{x}{\sigma^2}\right) = -\frac{1}{\sigma^4} e^{x - \frac{x^2}{2\sigma^2}} \left(x^2 - \sigma^2 x - \sigma^2\right).$$

By solving $x^2 - \sigma^2 x - \sigma^2 = 0$, we find that $h(x)$ is maximized when $x = \frac{\sigma^2 + \sqrt{\sigma^4 + 4\sigma^2}}{2}$. Denote this value by $\hat{x}(\sigma^2)$. Then $M(\sigma^2) = \sup_{x \geq 0} \frac{f(x)}{g(x)} = \frac{\hat{x}(\sigma^2)}{\sigma^2} \exp\left(-\hat{x}(\sigma^2) - \frac{1}{2\sigma^2}\left(\hat{x}(\sigma^2)\right)^2\right)$. One would repeatedly sample $x \sim \mathsf{Exp}(1)$ and independently $u \sim \mathsf{Unif}(0,1)$ until $uM(\sigma^2) \leq r(x)$, and then accept x in that case. The probability of acceptance of a random sample x is then $1/M(\sigma^2)$.

6. (1 pts) von Mises distribution is a continuous probability distribution on a circle with pdf

$$f(x|\mu,\kappa) \propto e^{\kappa \cos(x-\mu)},\ x \in [-\pi, \pi)$$

with parameters $\mu \in \mathbb{R}$ and $\kappa > 0$. Propose an algorithms to draw samples according to f.

Solution: This problem also has multiple solutions. We solve it using Accept-Reject method. Set $\hat{f}(x) = e^{\kappa \cos(x-\mu)}$. Note that $\hat{f}(x) \leq e^{\kappa}$ for all $x \in [-\pi, \pi)$. We will use uniform on $[-\pi, \pi)$, $g(x) = \frac{1}{2\pi}$ as the proposal density. Then

$$\frac{f(x)}{g(x)} = \frac{2\pi \hat{f}(x)}{C} \leq \frac{2\pi e^{\kappa}}{C},$$

so we can set $MC = 2\pi e^{\kappa}$ (ensuring that $\frac{f(x)}{g(x)} \leq M$). To draw samples from f, we repeatedly draw $x \sim g = \mathsf{Unif}[-\pi, \pi)$ and (independently) $u \sim \mathsf{Unif}(0,1)$ until $2\pi e^{\kappa} u \leq \frac{\hat{f}(x)}{g(x)}$ (equivalent to $\frac{f(x)}{Mg(x)} \leq u$) upon which we accept x and stop.

7. (2 pts) Gumbel's bivariate exponential distribution has cdf

$$F(x, y; k) = 1 - e^{-x} - e^{-y} + e^{-x-y-kxy}, x, y > 0$$

where $k \in [0,1]$ is a parameter. Describe how to obtain bivariate samples from F.

Solution: We will sample $(x, y) \sim F$ by first sampling $x \sim F_X(x)$, and then $y|x \sim F_{Y|X}(y|x)$. Note that $F_X(x) = \lim_{y \to \infty} F(x, y) = 1 - e^{-x}$; thus both marginals are univariate standard exponentials. Thus one can draw $x \sim \mathsf{Exp}(1)$. For absolutely continuous random variables x and y,

$$F_{y|x}(y|x) = \frac{1}{f_X(x)} \frac{\partial F(x,y)}{\partial x}.$$

Thus

$$F_{Y|X}(y|x) = e^{x}\left(e^{-x} - (1 + ky)\, e^{-x-y-kxy}\right) = 1 - (1 + ky)\, e^{-(1+kx)y},\ y > 0.$$

While $F_{Y|X}(y|x)$ cannot be inverted analytically, one can find y corresponding to $u \sim \mathsf{Unif}(0,1)$ using numerical methods (e.g., bisection).