STAT 535

EVALUATING DISTRIBUTION FUNCTIONS AND QUANTILE FUNCTIONS

dnorm(), rnorm(), pnorm(), qnorm()

dbeta(), rbeta(), pbeta(), qbeta()

dgamma(), rgamma(), pgamma(), qgamma()

etc, etc.

We take for granted - *what's under the hood?*

Power Series

Example: Beta cdf.

$$F(x) = \int_0^x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} z^{a-1}(1-z)^{b-1}dz$$

Can (will) show that

$$F(x) = x^a(1-x)^b \sum_{i=0}^{\infty} c_i x^i$$

where

$$c_0 = \frac{\Gamma(a+b)}{\Gamma(a+1)\Gamma(b)}$$

$$c_i = \left(\frac{i-1+a+b}{i+a}\right) c_{i-1}$$

Converges fast enough to yield a useful approximation?


```
mypbeta1 <- function(x, a, b, n=20) {
  cf <- exp(lgamma(a+b)-log(a)-lgamma(a)-lgamma(b))
  sum <- cf
  for (i in 1:n) {
    cf <- cf * (i-1+a+b)/(i+a)
    sum <-sum + cf*x^i
    print(c(i, x^a*(1-x)^b*sum), digits=5) } }
```

```
> mypbeta1(.2, 5, 15)
[1]  1  0.21820
[1]  3  0.29937
[1]  5  0.32116
[1]  7  0.32576
[1]  9  0.32657
[1] 11  0.32669
[1] 13  0.32671
```

```
> round(pbeta(.2, 5, 15)
[1] 0.32671
```


Art and Ad-Hockery...

Consider two routes to evaluating the standard normal cdf

$$\Phi(x) = 0.5 + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-y^2/2}dy$$

Can do power series expansion of $e^{-y^2/2}$ about $y = 0$ and integrate term-by-term. Yields alternating series expansion for $\Phi(x)$.

Alternative

$$\Phi(x) = 0.5 + \frac{1}{\sqrt{2\pi}} e^{-x^2/2} g(x)$$

where $g(x) = e^{x^2/2} \int_0^x e^{-y^2/2}dy$. Can work out expansion (how?) for $g(x)$ directly.

Comparison: when x=qnorm(.75)

```
# terms       A         B
     2   0.74868  0.74684
     3   0.75007  0.74980
     4   0.75000  0.74999
     5   0.75000  0.75000
```

But when x=4 [round(pnorm(4),5)=0.99997]

```
# terms       A         B
     2  -2.15962  0.50339
     3   8.05331  0.51253
     4 -11.39988  0.53341
     5  18.86063  0.57053

    11  10.1518   0.9276
    12  -4.77881  0.95738
```



```
    13   4.37861  0.97644
    14  -0.83929  0.98774
    15   1.93673  0.99397

    21   1.00493  0.99994
    22   0.99823  0.99996
    23   1.00055  0.99997
    24   0.99978  0.99997
    25   1.00003  0.99997
```


Continued Fraction Expansions

Gauss (1812) for hypergeometric functions, connection between power series expansions and continued fraction expansions

Example - Beta(a,b) cdf can be expressed as

$$F(x) \quad = \quad f(x)g(x)$$

where $f(x)$ is closed-form function (close to corresponding density function) and

$$g(x) \quad = \quad \frac{1}{1+}\,\frac{d_1 x}{1+}\,\frac{d_2 x}{1+}\ldots$$

where d_i has a simple form - depends on i, a, b.

What does this notation really mean?

Conventional wisdom: continued fraction expansions often converge faster than power series expansions, and may have wider ROC.


```
mypbeta2 <- function(x, a, b, n) {
  cf <- rep(0,n)
  for (i in (0:((n/2)-1))) {
    cf[2*i+1] <- -(a+b+i)*(a+i)/(a+2*i)/(a+2*i+1)
    cf[2*i+2] <- -(i+1)*(i+1-b)/(a+2*i+1)/(a+2*i+2) }
  ans <- 1+cf[n]*x
  for (i in ((n-1):1)) { ans <- 1 + cf[i]*x / ans }
  ans <- 1/ans
  ans <- exp(lgamma(a+b)-lgamma(a+1)-lgamma(b))*
       x^a*(1-x)^b*ans
  ans }

> round(mypbeta2(.2, 5, 15, n=6), digits=5)
[1] 0.32674
> round(mypbeta2(.2, 5, 15, n=8), digits=5)
[1] 0.32671
```

What about Quantiles

$$Q(\alpha) = F^{-1}(\alpha)$$

i.e., need to solve

$$F(x) = \alpha$$

for x, where we have some numerical routine to evaluate (approximate) $F()$.

In the absence of anything more sophisticated, one could use *bisection*.

If the mean and SD of F() are available, they might be useful for finding an *intitial bracketing*.