USING MAPLE FOR INSTRUCTION IN UNDERGRADUATE PROBABILITY AND STATISTICS

Elliot A. Tanis, Hope College, USA

Computers have been used in a variety of ways to enhance instruction in probability and statistics. A Computer Algebra System (CAS) such as Maple opens up new opportunities for developing insights. A CAS can provide a deeper understanding of basic concepts. For example a CAS can help students understand more completely various probability models and the effect of changes in their parameters. Figures that are static in a book can come alive using animation. Problems that are difficult or tedious to solve "by hand" can be solved quite easily using a CAS, sometimes symbolically. Random phenomena can be simulated and the data compared numerically and graphically with a theoretical model. Illustrative examples are given.

INTRODUCTION

There are a variety of problems that can be solved using a computer algebra system (CAS). By using a CAS in the classroom, fundamental principles of probability and statistics can be communicated in an active teaching/learning environment.

A CAS can make certain calculations easier. A CAS can be useful for finding the sum of particular infinite series. At times integrals must be evaluated. A CAS can replace books of integrals. Animation can be an effective tool to show how changes in parameters affect the shape of probability density functions or how the graphs of power functions change for different sample sizes or different significance levels. Distributions of functions of random variables, for example, convolutions, can be solved using a CAS. It is possible to provide a deeper understanding of the distribution of order statistics using a CAS.

There is a statistics package that comes with *Maple*. However it is not complete for all applications. Zaven Karian from Denison University in Ohio took the lead in writing more than 100 procedures to supplement *Maple*(Karen, 1994). These procedures always begin with a capital letter and sometimes have additional capital letters within their name. Almost all of *Maple*'s procedures use lower case letters exclusively.

Instead of discussing applications abstractly, a variety of illustrative examples are given. For the examples, the problem is stated and then *Maple* commands are given for solving the problem. Because of space limitations, only a few examples are given here. The author is willing to share additional examples.

SUMMATION EXAMPLES

(a) Let X equal the number of rolls of a pair of dice that are needed to determine whether a bettor wins or loses playing craps. The p.d.f. of X is $f(1) = \frac{12}{36}$ and

$$f(x) = 2\left[\left(\frac{3}{36}\right)\left(\frac{9}{36}\right)\left(\frac{27}{36}\right)^{(x-2)} + \left(\frac{4}{36}\right)\left(\frac{10}{36}\right)\left(\frac{26}{36}\right)^{(x-2)} + \left(\frac{5}{36}\right)\left(\frac{11}{36}\right)\left(\frac{25}{36}\right)^{(x-2)}\right]$$

for $x = 2,3,\ldots$. Show that $E(X) = \mu = \frac{557}{165}$ and $\mathrm{Var}(X) = \sigma^2 = \frac{245{,}672}{27{,}255}$. The *Maple* code that is needed to solve this problem is:

```
>f := 2*((27/36)^(x-2)*(3/36)*(9/36) +
        (26/36)^(x-2)*(4/36)*(10/36) + (25/36)^(x-2)*(5/36)*(11/36));
>mu := 12/36 + sum(x*f, x = 2 .. infinity);
>var := 12/36*(1 - mu)^2 + sum((x - mu)^2*f, x = 2 .. infinity);
```

(b) Let X have a Poisson distribution with mean $\mu = \lambda$. The following *Maple* code will find the mean and variance of X using summations and also the first three moments using the moment-generating function.

```
> f := PoissonPDF(lambda, x);  # Poisson p.d.f.
> mu := simplify(sum(x*f, x = 0 .. infinity));
> var := simplify(sum((x - mu)^2*f, x = 0 .. infinity));
> M := sum(exp(t*x)*f, x = 0 .. infinity);  # moment-generating function
> Mp := diff(M, t);
> mu := simplify(subs(t = 0, Mp));
> Mpp := diff(M, t$2);
> var := simplify(subs(t = 0, Mpp) - mu^2);
> Mppp := diff(M, t$3);
> ThirdM := simplify(subs(t = 0, Mppp));
```

INTEGRATION EXAMPLES

(a) Let X have a normal distribution with mean μ and variance σ^2. We can use the following *Maple* commands to find the moment-generating function of X.

```
> assume(sigma > 0);
> f := NormalPDF(mu, sigma^2, x);
>M := int(exp(t*x) *f, x = -infinity .. infinity);
```

(b) Let $f(x) = xe^x$, $0 < x < 1$. We use *Maple* to show that $f(x)$ is a p.d.f., $\mu = e - 2$, and $\sigma^2 = 2e - e^2 + 2$.

```
> f := x * exp(x);
> int(f, x = 0 .. 1);
> mu := int(x * f, x = 0 .. 1);
> var := int((x - mu)^2 * f, x = 0 .. 1);
```

ANIMATION EXAMPLES

(a) To illustrate the Central Limit Theorem for a discrete distribution, we can use animation to plot the probability histogram for the Poisson distribution for λ going from 1 to 25 with the $N(\lambda, \lambda)$ p.d.f. superimposed.

```
> for L from 7 to 25 do  # L = lambda
> H := ProbHist(PoissonPDF(L, x),  0 .. 40):
> N := plot(NormalPDF(L, L, x), x = 0 .. 40):
> t := textplot([30, 0.12, `lambda =` .L]):
> H.L := display({H, N, t}):
> od:
> display([seq(H.L,  L = 7 .. 25)], insequence = true);
```

(b) Animation can be an effective tool for illustrating how the sample size affects the power function. We shall use a test of hypotheses about the mean for a normal distribution. Assume that X is $N(\mu, 144)$. Test $H_0: \mu = 50$ against the alternative hypothesis $H_1: \mu > 50$. Fix the significance level at $\alpha = 0.05$. For each n = 8, 16, 24, 32, 40, 48, 56, and 64, graph the respective power functions using animation and then superimpose the eight power functions on one graph.

```
> zcrit := NormalP(0, 1, 0.95);
> for k from 1 to 8 do
>   n := 8 * k:
>   c := 50 + zcrit * 12/sqrt(n):
>   K := mu -> 1 - NormalCDF(mu, 144/n, c):
>   t := textplot([58, 0.064 * k, `n = `.n]):
```

```
>   P := plot(K(mu), mu = 50 .. 60, y = 0 .. 1):
> Q.k := display({P, t}):
> od:
> display([seq(Q.k, k = 1 .. 8)], insequence = true);
> display({seq(Q.k, k = 1 .. 8)});
```

DISTRIBUTION THEORY EXAMPLES

(a) It is not difficult to find the p.d.f. for the sum of the rolls of two fair 6-sided dice. It is more challenging to find the p.d.f. of the sum of the rolls of twelve 6-sided dice, but a convolution program can do this easily. We set up the probabilities of each die face in a list, use the Convolution program, and plot a probability histogram as follows.

```
> Y := [1, 1/6, 2, 1/6, 3, 1/6, 4, 1/6, 5, 1/6, 6, 1/6];
> X2 := Convolution(Y, Y); # sum of 2 dice
> X4 := Convolution(X2, X2): # sum of 4 dice
> X6 := Convolution(X4, X2): # sum of 6 dice
> X12 := Convolution(X6, X6): # sum of 12 dice
> P12 := ProbHist(X12):
> T12 := plot([[20,0], [64,0]], title = `Probability Histogram, Sum of 12 Dice`):
> display({P12, T12});
```

(b) A CAS can be used to show that the weighted average of two Cauchy random variables is Cauchy. This in turn can be used to prove that the distribution of $\overline{X}$ is Cauchy when sampling from the Cauchy distribution. That is, when sampling from a Cauchy distribution, the distribution of $\overline{X}$ does not become normal but is always the same as the parent distribution.

Let X_1 and X_2 have independent Cauchy distributions and let $X_3 = p/(p+q)*X_1 + q/(p+q)*X_2$ be their weighted average where $p > 0, q > 0$. If *f(x)* is the Cauchy p.d.f, then $p/(p+q)*X_1$ will have the p.d.f. $(p+q)/p*f(x*(p+q)/p))$. Similarly, $q/(p+q)*X_2$ will have the p.d.f. $(p+q)/q*f(x*(p+q)/q))$. The p.d.f. of X_3 is just the convolution of these two p.d.f.'s. In the expanded *Maple* package CauchyPDF(x) gives the Cauchy p.d.f. centered at 0. So the required *Maple* code for solving this problem is:

```
> assume(p > 0):
> additionally(q > 0):
> simplify(int(((p + q)/p* CauchyPDF(t* (p + q)/p)) *
   ((p + q)/q*  CauchyPDF((x - t) * (p + q)/q)), t = -infinity .. infinity));
```

Now, if $p = q = 1$, the weighted average just comes out to be the mean of the two random variables, X_1 and X_2, which implies that the mean of two independently distributed Cauchy random variables has a Cauchy distribution. By fixing *q* at 1 and using induction on *p*, it follows that the mean of any number of independently distributed Cauchy random variables has a Cauchy distribution.

ORDER STATISTICS EXAMPLE

(a) Define the p.d.f. for the *r*th order statistic in a random sample of size *n* for a given distribution function. Then use animation to graph the p.d.f.' s of the order statistics, $Y_1, Y_2, \ldots Y_n$, from *U*(0, 10), the uniform distribution on the interval (0, 10). For this example, the distribution function, $F(x) = x/10,\ 0 \le x \le 10$ is defined as an expression.

```
> OrderStatPDF := proc(expr::algebraic, n::posint, r::posint)
> local f;
> f := diff(expr,x):
> n!/(r-1)!/(n-r)! * expr^(r-1) * (1 - expr)^(n-r) * f;
> end:
> F := x/10;
> n := 9:
> for r from 1 to n do
> g[r] := OrderStatPDF(F, n, r);
> ex := int(x* g[r], x = 0..10): # mean of r'th order statistic
> t1 := textplot([5, 0.6, `r = `.r]):
> t2 := textplot([5, .4, `E(Y) = `.ex]):
> P := plot(g[r], x = 0 .. 10, title = `U(0, 10) order statistics, n = 9`):
> Q.r := display({P, t1, t2});
> od:
> display([seq(Q.k, k = 1 .. n)], insequence=true);
```

SUMMARY

The above examples give illustrations of applications of a CAS in a calculus-based probability and statistics course. The author has many more examples and is very willing to share them. Write to tanis@hope.edu. Also see Karian and Tanis (1995) for many exercises.

Karian from Denison University did the major development of the extra *Maple* procedures that are used. These are available free of charge as a *Maple* package for Release 4. For more information contact Karian@Denison.edu.

Many Hope College students have contributed to my understanding of *Maple* and provided applications in probability and statistics. Among them are Bryan Goodman, Joshua Levy, John Krueger, and Michael Van Opstall and I thank each of them for their contributions.

REFERENCES

Karian, Z. A. (1994). “Computer Algebra Systems in Statistics *Proceedings of the Fourth International Conference on Teaching Statistics*.

Karian, Z. A. and Tanis, E. A. (1995). *Probability and Statistics Explorations with MAPLE*, Prentice-Hall, NJ.