Moment Generating Functions

Definition and properties of MGF

For a probability density f[x] ≥ 0 with $\int_{-\infty}^{+\infty} f[x]\, dx = 1$ we define the MGF (moment generating function)

$$g[r] = \int_{-\infty}^{+\infty} e^{r x} f[x]\, dx$$

which exists uniquely for each real number r since the integrand is not negative. The integral may however take the value +∞ (i.e. the integral defining the moment generating function may for some r be divergent).

The same considerations apply to a discrete density p[x] ≥ 0 with p[x] > 0 only for enumerably many values of x and $\sum_x p[x] = 1$.

If the integral g[r] is finite for some value r1 < 0 and also for some other value r2 > 0 then it is known that g[r] is necessarily finite for every r between these values. It implies that g[r] will necessarily be finite on an interval neighborhood of the origin r = 0. In such a case all derivatives of g exist at r = 0 and may be computed under the integral! This means that in such a case (for every k) the k-th derivative at r = 0 exists and satisfies

$$\frac{d^k}{dr^k}\, g[r]\, |_{r=0} = \int_{-\infty}^{+\infty} \frac{d^k}{dr^k}\, e^{r x} f[x]\, dx$$

$$= \int_{-\infty}^{+\infty} x^k e^{0 x} f[x]\, dx = \text{E } X^k .$$

That is, we can obtain the integer moments of r.v. X !!!

Example: MGF for Bernoulli

Bernoulli r.v. X with P(X = 1) = p, P(X = 0) = q.   We have discrete probability density p[0] = q, p[1] = p and MGF

$$g[r] = p\, e^{r\,1} + q\, e^{r\,0} = p\, e^{r}$$

which is finite for every real number r. Hence g is finite for values of r in a neighborhood of 0.

Let us find E X froom theMGF. We have E X = g'[0] = p e^{0} = p.

Let us find E X^2 from the MGF. It is E X^2 = g''[0] = p e^{0} = p.

Therefore the variance Var X = E X^2 - $(E\,X)^2$ = p - p^2 = p q.

Example: MGF for Binomial

Binomial distribution n, p. We have r.v. S = $X_1 + ... + X_n$ where r.v. X_i are i.i.d. Bernoulli p. Because of independence we have

$$E\, e^{r S} = E\, e^{r (X_1 + ... + X_n)} = E\, e^{r X_1} ... E\, e^{r X_n}$$
$$= g[r] ... g[r] = (p\, e^r + q)^n$$

Let us find E S (which we know to be np) from the above MGF for S which is (different from Bernoulli) g[r] = $(p\, e^r + q)^n$.

$$E\, S = \frac{d}{dr} (p\, e^r + q)^n \text{ (at r = 0)}$$

$$= n\, (p\, e^r + q)^{n-1}\, p\, e^r \text{ (at r = 0)}$$

$$= n\, (p\, e^0 + q)^{n-1}\, p\, e^0 = np.$$

Let us find E S^2 from the MGF for r.v. S.

$$E\, S^2 = \frac{d^2}{d r^2} (p\, e^r + q)^n \text{ (at r = 0)}$$

$$= \frac{d}{d r}\, n\, (p\, e^r + q)^{n-1}\, p\, e^r \text{ (at r = 0}$$

$$= n\, p\, [\, (n-1)\, (p\, e^r + q)^{n-2} p\, e^r + (p\, e^r + q)^{n-1} e^r\,]$$

at r = 0 is

$$= n\, p\, [(n-1)\, p + 1]$$

So Var S (which we know to be npq) is by the MGF

$$\text{Var } S = E\, S^2 - (E\, S)^2 = n\, p\, [(n-1)\, p + 1] - (np)^2$$
$$= n\, p\, [\, -p + 1] = n\, p\, q.$$

Using *Mathematica* to evaluate the derivatives.

In[1]:= **g[x_] := (p Exp[r] + q) ^n**

In[2]:= **? D**

```
D[f, x] gives the partial derivative of f with respect to x. D[f, {x, n}] gives the
  nth partial derivative of f with respect to x. D[f, x1, x2, ... ] gives a mixed derivative.
```

In[3]:= **D[g[r], {r, 1}]**

Out[3]= $E^r\, n\, p\, (E^r p + q)^{-1+n}$

In[4]:= **D[g[r], {r, 2}]**

Out[4]= $E^{2r}\, (-1+n)\, n\, p^2\, (E^r p + q)^{-2+n} + E^r\, n\, p\, (E^r p + q)^{-1+n}$

In[5]:= **? Series**

```
Series[f, {x, x0, n}] generates a power series expansion for f about the point x =
  x0 to order (x - x0)^n. Series[f, {x, x0, nx}, {y, y0, ny}] successively finds
  series expansions with respect to y, then x.
```

```
Series[f, {x, x0, n}] generates a power
  series expansion for f about the
  point x = x0 to order (x - x0)^n.
```

In[6]:= **Series[g[r], {r, 0, 2}]**

Out[6]= $(p+q)^n + n\, p\, (p+q)^{-1+n}\, r + \left(\frac{1}{2}\, (-1+n)\, n\, p^2\, (p+q)^{-2+n} + \frac{1}{2}\, n\, p\, (p+q)^{-1+n}\right) r^2 + O[r]^3$

So we see from the series that the coefficient of "r" is

$$n\,p\,(p + q)^{n-1} = n\,p = E\,S \quad (\text{since } p + q = 1)$$

and also TWICE the coefficient of "r^2" is

$$(n\text{-}1)\,n\,p^2 + n\,p = np\,((n\text{-}1)\,p + 1)$$

which we got above. So Var S = npq by this Series method also.

MGF for the uniform on interval [a, b]

The MGF for the uniform on [a, b] is

$$g[r] = \frac{1}{b-a}\left(\frac{e^{b\,r}}{r} - \frac{e^{a\,r}}{r}\right)$$

We expect the derivatives of g to have indeterminate forms 0/0 which may confuse *Mathematica*.

```
In[7]:= g[r_] :=
          (Exp[ b r] - Exp [a r]) / ( r (b - a))
```

```
In[8]:= D[g[r], {r, 1}]
```

$$\text{Out[8]=}\; -\frac{-E^{a\,r} + E^{b\,r}}{(-a+b)\,r^2} + \frac{-a\,E^{a\,r} + b\,E^{b\,r}}{(-a+b)\,r}$$

```
In[9]:= D[g[r], {r, 2}]
```

$$\text{Out[9]=}\; \frac{2\,(-E^{a\,r} + E^{b\,r})}{(-a+b)\,r^3} - \frac{2\,(-a\,E^{a\,r} + b\,E^{b\,r})}{(-a+b)\,r^2} + \frac{-a^2\,E^{a\,r} + b^2\,E^{b\,r}}{(-a+b)\,r}$$

So the derivatives need careful examination to see what happens at r = 0. We resort to having *Mathematica* carry out a Taylor's Series expansion approximation of g[r] near r = 0.

```
In[10]:= Series[g[r], {r, 0, 2}]
```

$$\text{Out[10]=}\; 1 + \frac{\left(-\frac{a^2}{2} + \frac{b^2}{2}\right)\,r}{-a+b} + \frac{\left(-\frac{a^3}{6} + \frac{b^3}{6}\right)\,r^2}{-a+b} + O[r]^3$$

Now we can "read off" the mean and second moment as the series coefficient of "r" and twice that of " r^2 " repsectively. So

$$\text{E U} = \frac{\left(-\frac{a^2}{2} + \frac{b^2}{2}\right)}{-a+b} = \frac{a+b}{2}$$

$$\text{E } U^2 = 2\ \frac{\left(-\frac{a^3}{6} + \frac{b^3}{6}\right)}{-a+b} = (a^2 + a\,b + b^2)\,/\,3$$

From these we find that Var U = $(b - a)^2 / 12$.

Some Example Questions

1. For the standard normal distribution the MGF is $g[r] = e^{\frac{r^2}{2}}$ for every real number r. Find the mean, second moment, and variance.

Here is a solution using Mathematica to do the derivatives.

In[11]:= `g[r_] := Exp[ r^2 / 2]`

In[12]:= `D[g[r], {r, 1}]`

Out[12]= $E^{\frac{r^2}{2}}\ r$

In[13]:= `D[g[r], {r, 2}]`

Out[13]= $E^{\frac{r^2}{2}} + E^{\frac{r^2}{2}}\ r^2$

From these derivatives we can evaluate at r = 0 to obtain

$$\text{E Z} = g'[0] = e^0\ 0 = 0$$

$$\text{E } Z^2 = g''[0] = e^0 + e^0\ 0 = 1$$

so Var Z = $1 - 0^2 = 1$.

Using the Taylor's Series expansion for g[r] around r = 0 we can "read off" many moments.

In[17]:=
```
Series[g[r], {r, 0 , 12}]
```

Out[17]= $1+\frac{r^2}{2}+\frac{r^4}{8}+\frac{r^6}{48}+\frac{r^8}{384}+\frac{r^{10}}{3840}+\frac{r^{12}}{46080}+O[r]^{13}$

All of the odd moments are zero since the coefficients of odd powers of r are all zero. This is no surprise since the standard normal distribution is symmetric around r = 0. But we read off

$E\,Z^4 = 4!\,(\text{coeff of } r^4) = 4!\,/\,8 = 3$

$E\,Z^6 = 6!\,(\text{coeff of } r^6) = 6!\,/\,48 = 15$

(there is a simple pattern to these).

2. The MGF for the Poisson-λ distribution is

$$g[r] = e^{\lambda(e^r-1)}$$

for every real number r. We find that

$$g'[r] = e^{\lambda(e^r-1)}\,\lambda\,e^r$$

so the mean of the Poisson is $g'[0] = \lambda$.

The second moment is gotten from

$$g''[r] = e^{\lambda(e^r-1)}\, \lambda\, e^r + e^{\lambda(e^r-1)}\, (\lambda e^r)^2$$

from which we find that the seond moment is g''[0] = $\lambda + \lambda^2$. Thus we find (for the Poisson) the variance to be the same as the mean λ.

Expanding g[r] in series we get

```
In[19]:= g[r_] := Exp[λ (Exp[r] - 1)]
In[21]:= Series[g[r], {r, 0, 2}]
```

Out[21]= $1 + \lambda r + \frac{1}{2}(\lambda + \lambda^2)\, r^2 + O[r]^3$

We read off the mean λ (the coeff of r) and the second moment ($\lambda + \lambda^2$) (twice the coeff of r^2). So the variance of the Poisson is also equal to the mean λ.

3. The MGF of the exponential (waiting time) distribution with parameter λ is

$$g[r] = \frac{\lambda}{\lambda - r}$$

for every real number r < λ. Since $\lambda > 0$ this means that we have all moments and

$$E\, T = g'[0] = \frac{1}{\lambda}$$

$$E\, T^2 = g''[0] = \frac{2}{\lambda^2}$$

So Var T = $\frac{1}{\lambda^2}$. We can also consult the series to look at the moments.

```
g[r_] := λ / (λ - r)
```

In[22]:=
```
Series[g[r], {r, 0, 2}]
```

Out[22]= $1 + \lambda r + \frac{1}{2} (\lambda + \lambda^2) r^2 + O[r]^3$

Notice that the computer program has not informed us that g[r] is infinite for r > λ.