# Moment generating functions and moments of random variables

Moment generating functions (MGFs) can be used to find moments of random variables, including the mean and variance. The $k$-th order moment of rv $X$, $\mathrm{E}(X^k)$, can be found by taking the $k$-th derivative of the MGF of $X$ and setting $t = 0$.

If $M_X(t)$ is the MGF of $X$, then $\mathrm{E}(X^k) = M_X^{(k)}(t)|_{t=0}$, where $M_X^{(k)}(t)$ is the $k$-th order derivative of $M_X(t)$.

Some differentiation rules you may find useful:

- If $f(t) = t^n$, then $f'(t) = nt^{n-1}$
- If $f(t) = e^{\alpha t}$, then $f'(t) = \alpha e^{\alpha t}$
- If $f(t) = g(h(t))$, then $f'(t) = g'(y)|_{y=f(t)} \times h'(t)$ (chain rule)

Example 1 (Binomial distribution): $M_X(t) = (q + pe^t)^n$ where $p + q = 1$. We can write $M_X(t) = g(h(t))$ with $g(y) = y^n$ and $h(t) = q + pe^t$. We find $g'(y) = ny^{n-1}$ and $h'(t) = pe^t$ and therefore $M_X'(t) = g'(y)_{y=h(t)=q+pe^t} \times h'(t) = n(q + pe^t)^{n-1} \times pe^t = pne^t(q + pe^t)^{n-1}$.

To find the moments of $X$ with the given MGF, you can take derivatives of $M_X(t)$ and set $t = 0$ or you can identify one of the distributions we considered in class. If you identify the distribution and its parameters correctly (in this example, binomial with parameters $n$ and $p$), then you can use the formula for the mean and variance for this distribution: $\mu = np$ and $\sigma^2 = npq = np(1 - p)$. No derivation is required in this case but you have to explain that MGF uniquely identifies the distribution and hence its moments.

Example 2 (Gamma distribution): $M_X(t) = 1/(1 - \lambda t)^n$. We can write $M_X(t) = g(h(t))$ with $g(y) = y^{-n}$ and $h(t) = 1 - \lambda t$. We find $g'(y) = -ny^{-n-1}$ and $h'(t) = -\lambda$ and therefore $M_X'(t) = g'(y)_{y=h(t)=1-\lambda t} \times h'(t) = -(1 - \lambda t)^{-n-1} \times (-\lambda) = n\lambda(1 - \lambda t)^{-n-1}$.

Similarly, we can find $M_X''(t) = n(n+1)\lambda^2(1-\lambda t)^{-n-2}$ and $M_X'''(t) = n(n+1)(n+2)\lambda^3(1-\lambda t)^{-n-3}$. It implies that $\mathrm{E}(X) = M_X'(t)|_{t=0} = n\lambda$, $\mathrm{E}(X^2) = M_X''(t)|_{t=0} = n(n+1)\lambda^2$ and $\mathrm{E}(X^3) = M_X'''(t)|_{t=0} = n(n+1)(n+2)\lambda^3$. The variance is $\mathrm{Var}(X) = \mathrm{E}(X^2) - \{\mathrm{E}(X)\}^2 = n(n+1)\lambda^2 - \{n\lambda\}^2 = n\lambda^2$.

Sum of geometric progression:

If $|q| < 1$, then $\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + \cdots = 1/(1 - q)$. You can use this result without proof. This property is useful to find moments and/or MGFs for some discrete distributions (such as the Poisson distribution we considered in class).

Example 3: Let $X$ be a random variable such that $\Pr(X = 1/2^k) = 1/2^k$, $k = 1, 2, 3, \ldots$. That is, $X$ takes values $1/2, 1/4, 1/8, \ldots$ with probabilities $1/2, 1/4, 1/8, \ldots$. Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ using the definition of the mean and variance.

Solution: To find $\mathrm{E}(X)$ and $\mathrm{E}(X^2)$, we use the above result about the sum of geometric progression with $q = 1/4$ and $q = 1/8$, respectively:

$$\mathrm{E}(X) = \sum_{k=1}^{\infty} \frac{1}{2^k} \Pr\left(X = \frac{1}{2^k}\right) = \sum_{k=1}^{\infty} \frac{1}{2^k} \times \frac{1}{2^k} = \sum_{k=1}^{\infty} \frac{1}{4^k} = \frac{1}{4} \sum_{k=0}^{\infty} \frac{1}{4^k} = \frac{1}{4} \times \frac{1}{1-1/4} = \frac{1}{3},$$

$$\mathrm{E}(X^2) = \sum_{k=1}^{\infty} \left(\frac{1}{2^k}\right)^2 \Pr\left(X = \frac{1}{2^k}\right) = \sum_{k=1}^{\infty} \frac{1}{4^k} \times \frac{1}{2^k} = \sum_{k=1}^{\infty} \frac{1}{8^k} = \frac{1}{8} \sum_{k=0}^{\infty} \frac{1}{8^k} = \frac{1}{8} \times \frac{1}{1-1/8} = \frac{1}{7},$$

$$\mathrm{Var(X)} = \mathrm{E}(X^2) - \{\mathrm{E}(X)\}^2 = \frac{1}{7} - \frac{1}{9} = \frac{2}{63}.$$

Example 4 (optional! Questions of this type will NOT be tested on the exam): Let $X$ be a random variable such that $\Pr(X = k) = 1/2^k$, $k = 1, 2, 3, \ldots$. That is, $X$ takes values $1, 2, 3, \ldots$ with probabilities $1/2, 1/2^2, 1/2^3, \ldots$. Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ using a) the definition of the mean and variance; b) the MGF of $X$.

Solution: a) We write:

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{k=1}^{\infty} k \Pr(X = k) = \sum_{k=1}^{\infty} \frac{k}{2^k} = \sum_{k=0}^{\infty} \frac{k+1}{2^{k+1}} = \sum_{k=0}^{\infty} \frac{k}{2^{k+1}} + \sum_{k=0}^{\infty} \frac{1}{2^{k+1}} \\
&= \frac{1}{2} \sum_{k=0}^{\infty} \frac{k}{2^k} + \frac{1}{2} \sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{2} \sum_{k=1}^{\infty} \frac{k}{2^k} + \frac{1}{2} \cdot \frac{1}{1-1/2} = \frac{1}{2}\mathrm{E}(X) + 1,
\end{aligned}$$

solving this equation we get $\mathrm{E}(X) = 2$.

$$\begin{aligned}
\mathrm{E}(X^2) &= \sum_{k=1}^{\infty} k^2 \Pr(X = k) = \sum_{k=1}^{\infty} \frac{k^2}{2^k} = \sum_{k=0}^{\infty} \frac{(k+1)^2}{2^{k+1}} = \sum_{k=0}^{\infty} \frac{k^2}{2^{k+1}} + \sum_{k=0}^{\infty} \frac{2k}{2^{k+1}} + \sum_{k=0}^{\infty} \frac{1}{2^{k+1}} \\
&= \frac{1}{2} \sum_{k=0}^{\infty} \frac{k^2}{2^k} + \sum_{k=0}^{\infty} \frac{k}{2^k} + \frac{1}{2} \sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{2} \sum_{k=1}^{\infty} \frac{k^2}{2^k} + \sum_{k=1}^{\infty} \frac{k}{2^k} + \frac{1}{2} \cdot \frac{1}{1-1/2} \\
&= \frac{1}{2}\mathrm{E}(X^2) + \mathrm{E}(X) + 1,
\end{aligned}$$

solving this equation we get $\mathrm{E}(X^2) = 2\mathrm{E}(X) + 2 = 6$. Finally, $\mathrm{Var}(X) = \mathrm{E}(X^2) - \{\mathrm{E}(X)\}^2 = 6 - 4 = 2$. b) We find the MGF of $X$, $M_X(t)$:

$$M_X(t) = \mathrm{E}(e^{Xt}) = \sum_{k=1}^{\infty} e^{kt} \Pr(X = k) = \sum_{k=1}^{\infty} e^{kt}/2^k = \sum_{k=1}^{\infty} (e^t/2)^k = \frac{e^t}{2} \sum_{k=0}^{\infty} (e^t/2)^k = \frac{e^t/2}{1 - e^t/2}.$$

We can recognize the MGF of the geometric distribution with $p = 1/2$.

Two more differentiation rules you may find useful:

- If $f(t) = g(t) \cdot h(t)$, then $f'(t) = g'(t) \cdot h(t) + g(t) \cdot h'(t)$

- If $f(t) = \frac{g(t)}{h(t)}$, then $f'(t) = \frac{g'(t) \cdot h(t) - g(t) \cdot h'(t)}{\{h(t)\}^2}$

We use the second rule first with $g(t) = 1$ and $h(t) = (1 - e^t/2)$ and then with $g(t) = e^t/2$ and $h(t) = (1 - e^t/2)^2$. To find $h'(t)$, chain rule can be used:

$$M_X(t) = -1 + \frac{1}{1 - e^t/2}, \quad M'_X(t) = \frac{e^t/2}{(1 - e^t/2)^2}, \quad M'_X(t)|_{t=0} = \frac{1/2}{(1 - 1/2)^2} = 2,$$

$$M''_X(t) = \frac{(e^t/2) \cdot (1 - e^t/2)^2 - (e^t/2) \cdot (-e^t/2) \cdot 2(1 - e^t/2)}{(1 - e^t/2)^4} = \frac{(e^t/2)(1 + e^t/2)}{(1 - e^t/2)^3},$$

$$\mathrm{E}(X^2) = M''_X(t)|_{t=0} = \frac{(1/2)(1 + 1/2)}{(1 - 1/2)^3} = 6, \quad \mathrm{Var}(X) = \mathrm{E}(X^2) - \{\mathrm{E}(X)\}^2 = 6 - 4 = 2.$$

If we recognize the geometric distribution from the MGF, we could avoid this derivation and use the results we proved in class: $\mathrm{E}(X) = 1/p = 1/(1/2) = 2$ and $\mathrm{Var}(X) = (1-p)/p^2 = (1 - 1/2)/(1/2)^2 = 2$. Again, in this case you have to explain that MGF uniquely identifies the distribution and hence its moments.