Moment Generating Function

The moment generating function $M(t)$ of the random variable X is defined for all real values t by

$M(t) = E[e^{tX}]$

$= \sum_x e^{tx} \Pr(X = x)$ if X is a discrete r.v.

All moments of X can be obtained from $M(t)$ and then evaluating the result at $t = 0$.

$M'(t) = \frac{d}{dt}E[e^{tX}] = E[\frac{d}{dt}(e^{tX})] = E[Xe^{tX}]$

$M'(0) = E[X]$

The nth derivative of $M(t)$ is given by

$M^n(t) = E[X^n e^{tX}] \ n \geq 1$

$M^n(0) = E[X^n] \ n \geq 1$

Example

Binomial Distribution $B(n, p)$:

$M(t) = E[e^{tX}] = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k}$

$= (pe^t + 1 - p)^n$

Chernoff Bounds

Let X be a r.v. and M be its MGF.

$\Pr(X \geq a) \leq e^{-ta}M(t)$ for all $t > 0$

$\Pr(X \leq a) \leq e^{-ta}M(t)$ for all $t < 0$

We obtain the best bound by using the t that minimizes RHS.

Chernoff Bound for Sum of Indicator R.V's

Theorem 1. *Let* $X_1, X_2, \ldots, X_n$ *be independent indicator random variables such that, for* $1 \leq i \leq n$, $\Pr[X_i = 1] = p_i$, *where* $0 < p_i < 1$. *Then, for* $X = \sum_{i=1}^{n} X_i$, $\mu = E[X] = \sum_{i=1}^{n} p_i$, *and any* $\delta > 0$,

$$\Pr\left(X > (1+\delta)\mu\right) < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

For $0 < \delta < 1$,

$$\Pr\left(X > (1+\delta)\mu\right) \leq e^{-\mu\delta^2/3}$$

For $R \geq 6\mu$

$$\Pr(X \geq R) \leq 2^{-R}$$

For $0 < \delta < 1$,

$$\Pr(X < (1-\delta)\mu) < e^{-\mu\delta^2/2}$$

Proof. Upper tail: For any positive real t,

$$\Pr\left(X > (1+\delta)\mu\right) = \Pr(e^{tX} > e^{t(1+\delta)\mu})$$

By Markov's inequality,

$$\Pr\left(X > (1+\delta)\mu\right) < \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}}$$

$$\begin{aligned}
&= \frac{E[e^{t\sum_{i=1}^{n} X_i}]}{e^{t(1+\delta)\mu}} = \frac{E[\Pi_{i=1}^{n} e^{tX_i}]}{e^{t(1+\delta)\mu}} = \frac{\Pi_{i=1}^{n} E[e^{tX_i}]}{e^{t(1+\delta)\mu}} \\
&= \frac{\Pi_{i=1}^{n}(p_i e^t + 1 - p_i)}{e^{t(1+\delta)\mu)}} = \frac{\Pi_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\
&< \frac{\Pi_{i=1}^{n} e^{p_i(e^t - 1)}}{e^{t(1+\delta)\mu}} = \frac{e^{\sum_{i=1}^{n} p_i(e^t - 1)}}{e^{t(1+\delta)\mu}} = \frac{e^{(e^t - 1)\mu}}{e^{t(1+\delta)\mu}} \\
&\le \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}
\end{aligned}$$

for $t = \ln(1+\delta)$

Using $\delta - (1+\delta)\ln(1+\delta) \le -\delta^2/3$ for $0 < \delta < 1$ we get

$$\Pr\left(X > (1+\delta)\mu\right) \le e^{-\mu\delta^2/3}$$

For $R \geq 6\mu$, $\delta \geq 5$.

$$\begin{aligned} Pr(X \geq (1+\delta)\mu) &\leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \\ &\leq \left(\frac{e}{6}\right)^{R} \\ &\leq 2^{-R}. \end{aligned}$$

Lower tail:

$$\Pr\left(X < (1-\delta)\mu\right) = \Pr\left(e^{-tX} > e^{-t(1-\delta)\mu}\right)$$

By Markov's inequality,

$$\Pr\left(X < (1-\delta)\mu\right) < \frac{E[e^{-tX}]}{e^{-t(1-\delta)\mu}}$$

Similar calculations yield

$$< \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}}$$

For $t = \ln(1/(1-\delta))$

$$\leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}} \right)^{\mu}$$

Since $(1-\delta)^{(1-\delta)} > e^{-\delta+\delta^2/2}$ we have

$$\Pr(X < (1-\delta)\mu) < e^{-\mu\delta^2/2}$$

$\square$