Lecture 23: Binomial exponential tail bounds

Lemma 23.1. *Let $X \sim \mathrm{Bi}(m,q)$ with $m \in \mathbb{N}$ and $0<q<1$.*

1. *If $0 < x \le m(1-q)$,*

$$\mathbb{P}(X - mq \ge x) \le \exp\left\{-m\left[\left(1-\tfrac{x}{m}-q\right)\log\left(\frac{1-\frac{x}{m}-q}{1-q}\right)+\left(\tfrac{x}{m}+q\right)\log\left(\frac{\frac{x}{m}+q}{q}\right)\right]\right\}. \quad (23.1)$$

2. *If $-mq \le x < 0$,*

$$\mathbb{P}(X - mq \le x) \le \exp\left\{-m\left[\left(1-\tfrac{x}{m}-q\right)\log\left(\frac{1-\frac{x}{m}-q}{1-q}\right)+\left(\tfrac{x}{m}+q\right)\log\left(\frac{\frac{x}{m}+q}{q}\right)\right]\right\}.$$

Proof. If $X \sim \mathrm{Bi}(m,q)$, then for any $\theta > 0$,

$$\mathbb{E}[e^{\theta X}] = \big((1-q)+qe^{\theta}\big)^m,$$

and so Lemma 16.2 from Lecture 16 implies that for $x > 0$,

$$\mathbb{P}(X - mq \ge x) \le \exp\big\{m\log\big((1-q)+qe^{\theta}\big) - \theta(x+mq)\big\}.$$

The exponent is minimised at

$$\theta = \log\left(\frac{(1-q)(\frac{x}{m}+q)}{q(1-\frac{x}{m}-q)}\right),$$

which is positive if $x < m(1-q)$. Thus we find that for $0 < x < m(1-q)$,

$$\mathbb{P}(X - mq \ge x) \le \exp\left\{-m\left[\left(1-\tfrac{x}{m}-q\right)\log\left(\frac{1-\frac{x}{m}-q}{1-q}\right)+\left(\tfrac{x}{m}+q\right)\log\left(\frac{\frac{x}{m}+q}{q}\right)\right]\right\}.$$

If $x = m(1-q)$, then the bound still holds by interpreting $0\log(0) = 0$, since both sides are then equal to q^m.

For the other tail, we apply the bound (23.1) to $(m - X) \sim \mathrm{Bi}(m, 1-q)$ to find, for $-mq < x < 0$,

$$\begin{aligned}\mathbb{P}(X - mq \le x) &= \mathbb{P}\big((m-X) - m(1-q) \ge -x\big)\\ &\le \exp\left\{-m\left[\left(1-\tfrac{x}{m}-q\right)\log\left(\frac{1-\frac{x}{m}-q}{1-q}\right)+\left(\tfrac{x}{m}+q\right)\log\left(\frac{\frac{x}{m}+q}{q}\right)\right]\right\}.\end{aligned}$$

Again at $x = -mq$, the bound holds by direct calculation. □

Remark 23.2 (Binomial tail bounds)**.** Similar to our discussion around Poisson tail probabilities in Lecture 16, we can get some understanding of the bound (23.1) by fixing q and setting $x = z\sqrt{mq(1-q)}$ for a fixed $z > 0$ and m as large. Then Taylor expansion implies $\log(1+a) = a - a^2/2 + \mathrm{O}(a^3)$ as $a \to 0$, which implies

$$\begin{aligned}\log\left(\frac{1-\frac{x}{m}-q}{1-q}\right) &= \log\left(1 - z\sqrt{\frac{q}{m(1-q)}}\right) = -z\sqrt{\frac{q}{m(1-q)}} - \frac{z^2}{2}\left(\frac{q}{m(1-q)}\right) + \mathrm{O}\big(m^{-3/2}\big),\\ \log\left(\frac{\frac{x}{m}+q}{q}\right) &= \log\left(1 + z\sqrt{\frac{1-q}{mq}}\right) = z\sqrt{\frac{1-q}{mq}} - \frac{z^2}{2}\left(\frac{1-q}{mq}\right) + \mathrm{O}\big(m^{-3/2}\big),\end{aligned}$$

and so some simplification yields

$$\left(1-\tfrac{x}{m}-q\right)\log\left(\frac{1-\frac{x}{m}-q}{1-q}\right)+\left(\tfrac{x}{m}+q\right)\log\left(\frac{\frac{x}{m}+q}{q}\right) = \frac{z^2}{2m} + \mathrm{O}\big(m^{-3/2}\big),$$

and so

$$\mathbb{P}\left(\frac{X-mq}{\sqrt{mq(1-q)}} \ge z\right) \le \exp\Big\{-z^2/2 + \mathrm{O}\big(m^{-1/2}\big)\Big\}; \quad (23.2)$$

compare to (16.7) from Lecture 16. Note that (23.2) is only an asymptotic bound (the bigo depends on z and p) and (23.1) should be used in applications.

Lemma 23.3 (Roch, Exercise 2.6)**.** *Let $X \sim \mathrm{Bi}(m, q)$ with $m \in \mathbb{N}$ and $0 < q < 1$, and $\mu := \mathbb{E}[X] = mq$. For any $0 < \delta < 1$,*

$$\mathbb{P}\big(X \geq (1+\delta)\mu\big) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \leq \exp\{-\mu\delta^2/3\}, \tag{23.3}$$

$$\mathbb{P}\big(X \leq (1-\delta)\mu\big) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \leq \exp\{-\mu\delta^2/2\}. \tag{23.4}$$

Proof. For the first inequality of (23.3), first note that the left hand side is zero if $\delta > (1-q)/q$, and hence the inequality is trivially satisfied, so we assume that $\delta \leq (1-q)/q$. Now use (23.1) with $x = \delta\mu \leq m(1-q)$, so that $x/m + q = (1+\delta)q$, to find

$$\begin{aligned}\mathbb{P}\big(X \geq (1+\delta)\mu\big) &\leq \exp\left\{-m\left[\big(1-(1+\delta)q\big)\log\left(\frac{1-(1+\delta)q}{1-q}\right) + (1+\delta)q\log(1+\delta)\right]\right\} \\ &= \left(\frac{\exp\left\{-\left(\frac{1-(1+\delta)q}{q}\right)\log\left(\frac{1-(1+\delta)q}{1-q}\right)\right\}}{(1+\delta)^{(1+\delta)}}\right)^{\mu},\end{aligned}$$

and so the first inequality of (23.3) is the same as

$$-\left(\frac{1-(1+\delta)q}{q}\right)\log\left(\frac{1-(1+\delta)q}{1-q}\right) \leq \delta.$$

To show this inequality, we write the left hand side as

$$-\left(\frac{1-(1+\delta)q}{q}\right)\log\left(\frac{1-(1+\delta)q}{1-q}\right) = -\left(\frac{1-q}{q}\right)\left(1-\frac{\delta q}{1-q}\right)\log\left(1-\frac{\delta q}{1-q}\right),$$

and so for $y := \delta q/(1-q) \in (0,1]$, we need to show

$$(1-y)\log(1-y) \geq -y. \tag{23.5}$$

If $y = 1$, then as usual interpreting $0\log(0) = 0$, the inequality is satisfied. For the remaining, the derivative of $\varphi(y) = (1-y)\log(1-y) + y$ is $\varphi'(y) = -\log(1-y)$, which is positive for $y \in (0,1)$, negative for $y \in (-1,0)$, and zero at $y = 0$. Moreover, the the second derivative is $\varphi''(y) = 1/(1-y)$, which is positive for $y < 1$, and so the minimum of φ is at 0 and is equal to 0. Thus $\varphi(y) \geq 0$, which is the same as (23.5).

The first inequality of (23.4) follows in almost exactly the same way after replacing δ with $-\delta$, and noting that we have shown (23.5) for $y \in (-1,1)$.

The second inequality of (23.3) is the same as

$$\delta - (1+\delta)\log(1+\delta) \leq -\delta^2/3. \tag{23.6}$$

Expanding in Taylor series, we have for $\delta \in (0,1)$,

$$\log(1+\delta) = \sum_{j=1}^{\infty}(-1)^{j-1}\frac{\delta^j}{j},$$

and plugging this in to the left hand side of (23.6) yields

$$\begin{aligned}\delta - (1+\delta)\log(1+\delta) &= \delta - \sum_{j=1}^{\infty}(-1)^{j-1}\frac{\delta^j}{j} - \sum_{j=1}^{\infty}(-1)^{j-1}\frac{\delta^{j+1}}{j} \\ &= -\sum_{j=2}^{\infty}(-1)^{j-1}\frac{\delta^j}{j} - \sum_{j=2}^{\infty}(-1)^{j}\frac{\delta^j}{j-1}\end{aligned}$$

$$= \sum_{j=2}^{\infty} (-1)^{j-1} \frac{\delta^j}{j(j-1)}$$
$$\leq -\frac{\delta^2}{2} + \frac{\delta^3}{6} \leq -\frac{\delta^2}{3},$$

where the first inequality is because the power series is alternating and the terms are decreasing in absolute value, and the second is because $\delta \in (0, 1)$.

The second inequality of (23.4) is the same as

$$-\delta - (1-\delta)\log(1-\delta) \leq -\delta^2/2,$$

which can be shown very similarly to (23.6) starting from

$$\log(1-\delta) = -\sum_{j=1}^{\infty} \frac{\delta^j}{j}.$$ □