Lecture10

Thursday, October 11, 2018 9:25 AM

Concentration of measures

* A few mathematical facts

<u>Lemma</u> (Markov inequality)

- Let m be a non-negative random variable. Then it holds:

$$\mathbb{P}[m \geq t] \leq \frac{\mathbb{E}[m]}{t} \quad \forall t > 0$$

<u>Proof</u>: We have

$$\mathbb{E}[m] = \mathbb{E}[m \, \mathbb{1}_{\{m \geq t\}}] + \underbrace{\mathbb{E}[m \, \mathbb{1}_{\{m < t\}}]}_{\geq 0}$$

$$\geq \mathbb{E}[m \, \mathbb{1}_{\{m \geq t\}}]$$

$$\geq t \, \mathbb{E}[\mathbb{1}_{\{m \geq t\}}] \qquad \int_S \mathbb{1}_{\{m \geq t\}} dm$$

$$\mathbb{E}[\mathbb{1}_{\{m \geq t\}}] = \mathbb{P}[m \geq t]$$

$$= t \, \mathbb{P}[m \geq t]$$

<u>Def</u> (moment generating functions) (MGF)

- for any random variable m, the following

$$\mathbb{E}[e^{\lambda m}] \qquad \forall \lambda$$

$$\mathbb{E}[e^{\lambda m}] \quad \forall \lambda$$

is called MGF of m

Ex (in your homework). For $m \sim N(\bar{m}, \sigma^2)$

Then:

$$\mathbb{E}[e^{\lambda m}] = e^{\lambda \bar{m} + \lambda^2 \sigma^2 / 2} \quad \forall \lambda \in \mathbb{R}$$

---

in your homework:

(Chebyshev) $\mathbb{P}[|m| \geq t] \leq \dfrac{\sigma^2}{t^2}$; $\sigma^2 = \mathbb{V}[m]$

(Chernoff) $\mathbb{P}[m > t] \leq \min\limits_{\lambda > 0} \dfrac{\mathbb{E}[e^{\lambda m}]}{e^{\lambda t}} \quad \forall t > 0$

In your homework:

$$\mathbb{P}[m - \bar{m} \geq t] \leq e^{-\frac{t^2}{2\sigma^2}} \quad \forall t \geq 0$$

$$m \sim N(\bar{m}, \sigma^2)$$

- This is called the tail bound of a random variable. → question is whether we can get a similar exponential tail bound for a more general class of random variables?

Proposition (12.1) (connection between tail bound and

Proposition (12.1) [connection between tail bound and expectation]. For <u>any</u> random variable m there holds:

$$\mathbb{E}[m] = \int_0^\infty \mathbb{P}[m > t]\,dt - \int_{-\infty}^0 \mathbb{P}[m < t]\,dt$$

<u>Proof</u>: assume $m \geq 0$; then we have

$$m = \int_0^m dt = \int_0^\infty \mathbb{1}_{\{m > t\}}\,dt$$

$$\Downarrow$$

$$\mathbb{E}[m] = \mathbb{E}\int_0^\infty \mathbb{1}_{\{m > t\}}\,dt$$

Fubini $\quad \| \quad \mathbb{E}[\mathbb{1}_{\{m > t\}}] = \mathbb{P}[m > t]$

$$\mathbb{P}[m > t]\,dt$$

generally $\quad m = m\,\mathbb{1}_{\{m \geq 0\}} + m\,\mathbb{1}_{\{m < 0\}}$

Lemma 12.3 (Union Bound)

$$\mathbb{P}[A_1 \cup A_2 \cup \dots \cup A_n] \leq \sum_{i=1}^{n} \mathbb{P}[A_i] \quad \forall n \in \mathbb{N}$$

Concentration of sum of scalar RVs

- If $m_i \sim N(\bar{m}, \sigma^2)$, recall

$$S_N := \frac{1}{N} \sum_{i=1}^{N} m_i \overset{??}{\sim} N\left(\bar{m}, \sigma^2/N\right)$$

$\Downarrow$ Chebyshev

$$\mathbb{P}[|S_N - \bar{m}| \geq t] \leq \frac{1}{N} \frac{\sigma^2}{t^2}$$

$\Downarrow$

$$\boxed{\lim_{N \to \infty} \mathbb{P}[|S_N - \bar{m}| \geq t] = 0 \quad \forall t > 0}$$

a - version of weak law of large numbers.

- We in fact have stronger results for Gaussians:

$$\boxed{\mathbb{P}[|S_N - \bar{m}| \geq t] \leq 2e^{-Nt^2/2\sigma^2}}$$

$\Downarrow$ true because

1) $\{ |S_N(\omega) - \bar{m}| \geq t \} = \{ S_N - \bar{m} \leq -t \} \cup$

$\{ S_N - \bar{m} \geq t \}$

2) $\mathbb{P}[A \cup B] \leq \mathbb{P}[A] + \mathbb{P}[B]$

3) for Gaussian:

$$\mathbb{P}[S_N - \bar{m} \geq t] \leq e^{-N \frac{t^2}{2\sigma^2}}$$

$$\mathbb{P}[S_N - \bar{m} \leq -t] \leq e^{-N t^2/2\sigma^2}$$

Summary: the tail of the sum of Gaussian random variables decays exponentially in both t and N. In other words

$$|S_N - \bar{m}| < t$$

holds with probability $1 - 2e^{-N \frac{t^2}{2\sigma^2}}$

because

$$\Omega = \{ |S_N - \bar{m}| < t \} \cup \{ |S_N - \bar{m}| \geq t \}$$

$$\mathbb{P}[\Omega] \underset{\text{union bound}}{\leq} \mathbb{P}[|S_N - \bar{m}| < t] + \underbrace{\mathbb{P}[|S_N - \bar{m}| \geq t]}_{\leq 2e^{-Nt^2/2\sigma^2}}$$

$\|$

$$\| \qquad \text{union bound} \qquad 2e^{-Nt^2/2\sigma^2}$$

$$1$$

$$\Downarrow$$

$$(*) \quad \boxed{\mathbb{P}[|S_N - \bar{m}| < t] \geq 1 - 2e^{-Nt^2/2\sigma^2}}$$

How about non-Gaussians

$$m_i \not\sim N(\bar{m}, \sigma^2)$$

It turns out the the tail bound (Gaussian tail) of the form $(*)$ is true for a large class of RVs. And we are going study the concentration phenomena through the tail bound of the form

$$\boxed{\mathbb{P}[|m - \bar{m}| \geq t] \leq \text{small quantity}}$$

- In other words, we study non-asymptotic behavior of sum of random variables.

Sub-Gaussian RVs

Def: a RV m is called sub-gaussian if

its MGF is bound by that of a zero mean Gaussian RV with variance σ^2, i.e.,

$$\mathbb{E}[e^{\lambda m}] \leq e^{\lambda^2\sigma^2/2}$$