

## Probability review

- *Probability space*  $\Omega$  - a finite (or for us at most countable) set endowed with a measure  $p : \Omega \rightarrow \mathcal{R}$  satisfying:

$$\forall \omega \in \Omega; p(\omega) \geq 0$$

and

$$\sum_{\omega \in \Omega} p(\omega) = 1.$$

- *An event*  $\mathbf{A} \subseteq \Omega$  -  $\Pr[\mathbf{A}] = \sum_{\omega \in \mathbf{A}} p(\omega)$ .

- *Random variable*  $\mathbf{X} : \Omega \rightarrow \mathcal{R}$  -

*Example:* If  $\mathbf{X}$  is a random variable then for a fixed  $t, t' \in \mathcal{R}$ ,  $t \leq \mathbf{X} \leq t'$  and  $\mathbf{X} > t$  are probabilistic events.

- *Two events*  $\mathbf{A}$  *and*  $\mathbf{B}$  *are independent* -  $\Pr[\mathbf{A} \cap \mathbf{B}] = \Pr[\mathbf{A}] \cdot \Pr[\mathbf{B}]$ .
- *Conditional probability of*  $\mathbf{A}$  *given*  $\mathbf{B}$  -  $\Pr[\mathbf{A}|\mathbf{B}] = \Pr[\mathbf{A} \cap \mathbf{B}] / \Pr[\mathbf{B}]$ .

*Example:*  $\mathbf{A}$  and  $\mathbf{B}$  are independent iff  $\overline{\mathbf{A}}$  and  $\mathbf{B}$  are independent iff ... iff  $\Pr[\mathbf{A}|\mathbf{B}] = \Pr[\mathbf{A}]$ .

- *For a random variable*  $\mathbf{X}$  *and an event*  $\mathbf{A}$ ,  $\mathbf{X}$  *is independent of*  $\mathbf{A}$  - for all  $S \subseteq \mathcal{R}$ ,  $\Pr[\mathbf{X} \in S | \mathbf{A}] = \Pr[\mathbf{X} \in S]$ .

- *Two random variables*  $\mathbf{X}$  *and*  $\mathbf{Y}$  *are independent* - for all  $S, T \subseteq \mathcal{R}$ ,  $\mathbf{X} \in S$  and  $\mathbf{Y} \in T$  are independent events.

- *Events*  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$  *are mutually independent* - for all  $I \subseteq \{1, \dots, n\}$ ,

$$\Pr\left[\bigcap_{i \in I} A_i \cap \bigcap_{i \notin I} \overline{A_i}\right] = \prod_{i \in I} \Pr[A_i] \cdot \prod_{i \notin I} \Pr[\overline{A_i}].$$

- *Random variables*  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  *are mutually independent* - for all  $t_1, t_2, \dots, t_n \in \mathcal{R}$ , events  $\mathbf{X}_1 = t_1, \mathbf{X}_2 = t_2, \dots, \mathbf{X}_n = t_n$  are mutually independent.

- *Expectation of a random variable*  $\mathbf{X}$  -  $\mathbf{E}[\mathbf{X}] = \sum_{\omega \in \Omega} p(\omega) \mathbf{X}(\omega)$ .

Three easy claims:

*Claim: (Linearity of expectation)* For random variables  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$

$$\mathbf{E}[\mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_n] = \sum_{i=1}^n \mathbf{E}[\mathbf{X}_i].$$

*Claim:* For independent random variables  $\mathbf{X}$  and  $\mathbf{Y}$ ,  $\mathbf{E}[\mathbf{X} \cdot \mathbf{Y}] = \mathbf{E}[\mathbf{X}] \cdot \mathbf{E}[\mathbf{Y}]$ .

*Claim:* For a random variable  $\mathbf{X} : \Omega \rightarrow \mathcal{N}$ ,  $\mathbf{E}[\mathbf{X}] = \sum_{k=1}^{\infty} \mathbf{E}[\mathbf{X} \geq k]$ .

*Theorem: (Markov Inequality)* For a non-negative random variable  $\mathbf{X}$  and any  $t \in \mathcal{R}$

$$\Pr[\mathbf{X} \geq t] \leq \frac{\mathbf{E}[\mathbf{X}]}{t}.$$

*Proof:*  $\mathbf{E}[\mathbf{X}] = \sum_{\omega \in \Omega} p(\omega) \mathbf{X}(\omega) \geq \sum_{\omega \in \Omega, \mathbf{X}(\omega) \geq t} p(\omega) \mathbf{X}(\omega) \geq t \cdot \sum_{\omega \in \Omega, \mathbf{X}(\omega) \geq t} p(\omega) = t \cdot \Pr[\mathbf{X} \geq t]$ .

*Theorem: (Chernoff Bounds)* Let  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  be independent 0-1 random variables. Denote  $p_i = \Pr[\mathbf{X}_i = 1]$ , hence  $1 - p_i = \Pr[\mathbf{X}_i = 0]$ . Let  $\mathbf{X} = \sum_{i=1}^n \mathbf{X}_i$ . Denote  $\mu = \mathbf{E}[\mathbf{X}] = \sum_{i=1}^n p_i$ . For any  $0 < \delta < 1$  it holds

$$\Pr[\mathbf{X} \geq (1 + \delta)\mu] \leq \left[ \frac{e^\delta}{(1 + \delta)^{(1 + \delta)}} \right]^\mu$$

and

$$\Pr[\mathbf{X} \leq (1 - \delta)\mu] \leq e^{-\frac{1}{2}\mu\delta^2}.$$

*Proof:* For any real number  $t > 0$ ,

$$\begin{aligned} \Pr[\mathbf{X} \geq (1 + \delta)\mu] &= \Pr[t\mathbf{X} \geq t(1 + \delta)\mu] \\ &= \Pr[e^{t\mathbf{X}} \geq e^{t(1 + \delta)\mu}] \end{aligned}$$

where based on  $\mathbf{X}$  we define new random variables  $t\mathbf{X}$  and  $e^{t\mathbf{X}}$ . Notice,  $e^{t\mathbf{X}}$  is a non-negative random variable so one can apply the Markov inequality to obtain

$$\Pr[e^{t\mathbf{X}} \geq e^{t(1 + \delta)\mu}] \leq \frac{\mathbf{E}[e^{t\mathbf{X}}]}{e^{t(1 + \delta)\mu}}.$$

Since all  $\mathbf{X}_i$  are mutually independent, random variables  $e^{t\mathbf{X}_i}$  are also mutually independent so

$$\mathbf{E}[e^{t\mathbf{X}}] = \mathbf{E}[e^{t \sum_i \mathbf{X}_i}] = \prod_{i=1}^n \mathbf{E}[e^{t\mathbf{X}_i}].$$

We can evaluate  $\mathbf{E}[e^{t\mathbf{X}_i}]$

$$\mathbf{E}[e^{t\mathbf{X}_i}] = p_i e^t + (1 - p_i) \cdot 1 = 1 + p_i(e^t - 1) \leq e^{p_i(e^t - 1)}.$$

where in the last step we have used  $1 + x \leq e^x$  which holds for all  $x$ . (Look on the graph of functions  $1 + x$  and  $e^x$  and their derivatives in  $x = 0$ .) Thus

$$\begin{aligned} \mathbf{E}[e^{t\mathbf{X}}] &\leq \prod_{i=1}^n e^{p_i(e^t - 1)} \\ &= e^{\sum_{i=1}^n p_i(e^t - 1)} \\ &= e^{\mu(e^t - 1)} \end{aligned}$$

By choosing  $t = \ln(1 + \delta)$  and rearranging terms we obtain

$$\begin{aligned} \Pr[\mathbf{X} \geq (1 + \delta)\mu] &= \Pr[e^{t\mathbf{X}} \geq e^{t(1 + \delta)\mu}] \\ &\leq \frac{e^{\mu(e^t - 1)}}{e^{t(1 + \delta)\mu}} \\ &= \left[ \frac{e^\delta}{(1 + \delta)^{(1 + \delta)}} \right]^\mu \end{aligned}$$

That proves the first bound. The second bound is obtained in a similar way:

$$\begin{aligned}\Pr[\mathbf{X} \leq (1 - \delta)\mu] &= \Pr[-t\mathbf{X} \geq -t(1 - \delta)\mu] \\ &= \Pr[e^{-t\mathbf{X}} \geq e^{-t(1 - \delta)\mu}] \\ &\leq \frac{\mathbf{E}[e^{-t\mathbf{X}}]}{e^{-t(1 - \delta)\mu}}.\end{aligned}$$

Bounding  $\mathbf{E}[e^{-t\mathbf{X}}]$  as before gives

$$\mathbf{E}[e^{-tX}] \leq e^{\mu(e^{-t} - 1)}$$

By choosing  $t = -\ln(1 - \delta)$  and rearranging terms we obtain

$$\begin{aligned}\Pr[\mathbf{X} \leq (1 - \delta)\mu] &= \Pr[e^{-t\mathbf{X}} \geq e^{-t(1 - \delta)\mu}] \\ &\leq \frac{e^{\mu(e^{-t} - 1)}}{e^{-t(1 - \delta)\mu}} \\ &= \left[ \frac{e^{-\delta}}{(1 - \delta)^{(1 - \delta)}} \right]^\mu\end{aligned}$$

We use the well known expansion for  $0 < \delta < 1$

$$\ln(1 - \delta) = -\sum_{i=1}^{\infty} \frac{\delta^i}{i}$$

to obtain

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

Thus

$$(1 - \delta)^{(1 - \delta)} \geq e^{\frac{\delta^2}{2} - \delta}$$

Hence

$$\Pr[\mathbf{X} \leq (1 - \delta)\mu] \leq e^{-\frac{\delta^2}{2} + \delta - \delta} = e^{-\frac{\delta^2}{2}\mu}$$

□