# 18.440: Lecture 38

# Review: practice problems

Scott Sheffield

MIT

# Order statistics

- Let $X$ be a uniformly distributed random variable on $[-1, 1]$.
  - Compute the variance of $X^2$.
  - If $X_1, \ldots, X_n$ are independent copies of $X$, what is the probability density function for the smallest of the $X_i$

# Order statistics — answers

- $$\mathrm{Var}[X^2] = E[X^4] - (E[X^2])^2$$
  $$= \int_{-1}^{1} \frac{1}{2} x^4 dx - (\int_{-1}^{1} \frac{1}{2} x^2 dx)^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.$$
- Note that for $x \in [-1, 1]$ we have
  $$P\{X > x\} = \int_{x}^{1} \frac{1}{2} dx = \frac{1-x}{2}.$$
  If $x \in [-1, 1]$, then
  $$P\{\min\{X_1, \ldots, X_n\} > x\}$$
  $$= P\{X_1 > x, X_2 > x, \ldots, X_n > x\} = (\frac{1-x}{2})^n.$$
  So the density function is
  $$-\frac{\partial}{\partial x}(\frac{1-x}{2})^n = \frac{n}{2}(\frac{1-x}{2})^{n-1}.$$

# Moment generating functions

- Suppose that $X_i$ are independent copies of a random variable $X$. Let $M_X(t)$ be the moment generating function for $X$. Compute the moment generating function for the average $\sum_{i=1}^n X_i/n$ in terms of $M_X(t)$ and $n$.

# Moment generating functions — answers

- Write $Y = \sum_{i=1}^n X_i/n$. Then
$$M_Y(t) = E[e^{tY}] = E[e^{t\sum_{i=1}^n X_i/n}] = (M_X(t/n))^n.$$

# Entropy

- Suppose $X$ and $Y$ are independent random variables, each equal to 1 with probability $1/3$ and equal to 2 with probability $2/3$.
  - Compute the entropy $H(X)$.
  - Compute $H(X+Y)$.
  - Which is larger, $H(X+Y)$ or $H(X,Y)$? Would the answer to this question be the same for any discrete random variables $X$ and $Y$? Explain.

# Entropy — answers

- $H(X) = \frac{1}{3}(-\log\frac{1}{3}) + \frac{2}{3}(-\log\frac{2}{3})$.
- $H(X+Y) = \frac{1}{9}(-\log\frac{1}{9}) + \frac{4}{9}(-\log\frac{4}{9}) + \frac{4}{9}(-\log\frac{4}{9})$
- $H(X,Y)$ is larger, and we have $H(X,Y) \geq H(X+Y)$ for any $X$ and $Y$. To see why, write $a(x,y) = P\{X = x, Y = y\}$ and $b(x,y) = P\{X+Y = x+y\}$. Then $a(x,y) \leq b(x,y)$ for any $x$ and $y$, so
  $H(X,Y) = E[-\log a(x,y)] \geq E[-\log b(x,y)] = H(X+Y)$.

18.440 Probability and Random Variables
Spring 2014