Chebyshev's Theorem (from last time):

Stat 561
10-23-25

1

$$P[g(X) \geq r] \leq \frac{E[g(X)]}{r}$$

Chebyshev's Inequality

Use the theorem with $g(X) = (X-\mu)^2$
and $r = t^2\sigma^2$ with $t>0$

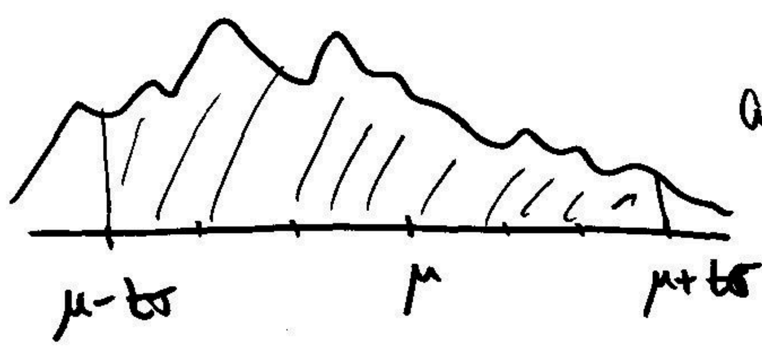
$$P[(X-\mu)^2 \geq t^2\sigma^2] \leq \frac{E[(X-\mu)^2]}{t^2\sigma^2} = \frac{\sigma^2}{t^2\sigma^2} = \frac{1}{t^2}$$

2

$$P[|X-\mu| \geq t\sigma] \leq \frac{1}{t^2}$$

$$\boxed{P[|X-\mu| < t\sigma] \geq 1 - \frac{1}{t^2}}$$

$P[A] \leq c$
$P[A^c] \geq 1-c$

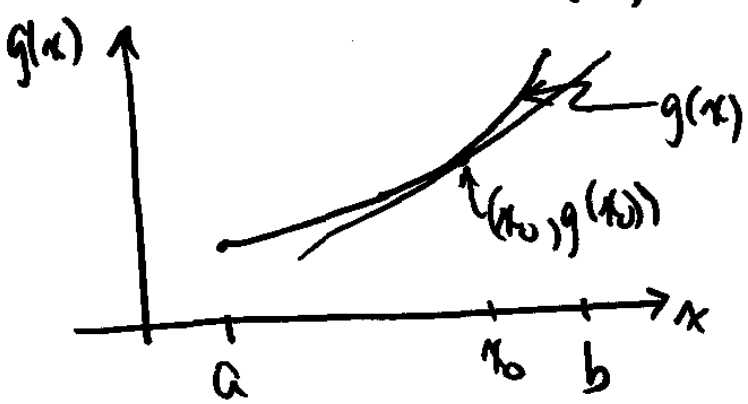


area $\geq 1 - \frac{1}{t^2}$

③

Defn: A function $g(x)$ defined on (a,b) is **convex** if $\forall$ $0<\gamma<1$
(strictly convex) $a<x_1<x_2<b$

$$g(\gamma x_1 + (1-\gamma)x_2) \leq \gamma g(x_1) + (1-\gamma) g(x_2)$$
$$(<)$$



④

Find the equation of the tangent line at x_0

The slope will be $g'(x_0)$

$$y - g(x_0) = g'(x_0)(x - x_0)$$

$$\underline{y = g(x_0) + g'(x_0)(x - x_0)}$$

Let's show that the curve lies above the tangent line.

(5)

$$g(\gamma x_1 + (1-\gamma)x_0) \le \gamma g(x_1) + (1-\gamma) g(x_0)$$
$$\le \gamma [g(x_1) - g(x_0)] + g(x_0)$$

$$\frac{g(\gamma x_1 + (1-\gamma)x_0) - g(x_0)}{\gamma} \le g(x_1) - g(x_0)$$

$$(x_1 - x_0)\frac{g(\gamma x_1 + (1-\gamma)x_0) - g(x_0)}{\gamma (x_1 - x_0)} \le g(x_1) - g(x_0)$$

Let $h = \gamma (x_1 - x_0)$ (6)

$$(x_1 - x_0)\frac{g(h + x_0) - g(x_0)}{h} \le g(x_1) - g(x_0)$$

Take limit as $h \to 0$

$$(x_1 - x_0)\, g'(x_0) \le g(x_1) - g(x_0)$$

$$g(x_1) \ge g(x_0) + g'(x_0)(x_1 - x_0)$$

Consequence: Suppose $g(x)$ is convex on (a,b)

$$E[g(X)] = \int_a^b g(x) f_X(x)\, dx$$

$$\geq \int_a^b [g(x_0) + g'(x_0)(x - x_0)] f_X(x)\, dx$$

$$= \int_a^b g(x_0) f_X(x)\, dx + \int_a^b g'(x_0)(x-x_0) f_X(x)\, dx$$

$$= g(x_0) + g'(x_0)\big[E[X] - x_0\big]$$

$\forall\ x_0 \in (a,b)$

8

In particular, this inequality holds for $x_0 = \mu$

$$\therefore\ E[g(X)] \geq g(\mu) + 0$$

That is, $E[g(X)] \geq g(E[X])$

<u>Jensen's Inequality</u>

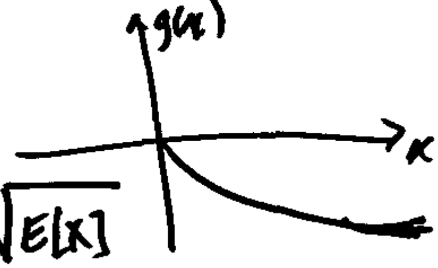
<u>Example</u>: $g(x) = x^2$ is convex

So $E[X^2] \geq (E[X])^2$

(9)

<u>Example</u>: $g(x) = -\sqrt{x}$ $\quad x > 0$

is convex



So $E[-\sqrt{X}] \geq -\sqrt{E[X]}$

$$E[\sqrt{X}] \leq \sqrt{E[X]}$$

<u>Example</u>: $g(x) = -\ln x$ $\quad x > 0$

is convex

$$E[-\ln X] \geq -\ln E[X]$$

$$E[\ln X] \leq \ln E[X]$$

(10)

Let X be a r.v. that takes on the values $a_1, a_2, \ldots, a_n$ with probability $\frac{1}{n}$ each.

$$E[X] = \sum_{i=1}^{n} x_i\, p(x_i) = \sum_{i=1}^{n} a_i \frac{1}{n} = \bar{a},$$

the arithmetic mean

$$E[\ln X] = \sum_{i=1}^{n} \ln(x_i)\, p(x_i) = \frac{1}{n} \sum_{i=1}^{n} \ln(a_i)$$

$$= \frac{1}{n} \ln\left(\prod_{i=1}^{n} a_i\right) = \ln \sqrt[n]{\prod_{i=1}^{n} a_i}$$

(11)

$$E[\ln X] \leq \ln E[X]$$

$$\ln \sqrt[n]{\prod_{i=1}^{n} a_i} \leq \ln(\overline{a})$$

$$\sqrt[n]{\prod_{i=1}^{n} a_i} \leq \overline{a}$$

geometric mean — arithmetic mean

(12)

Defn: The discrete uniform distribution

X takes on the values $\{1, 2, \ldots, N\}$

with probabilities $\frac{1}{N}$ each.

$$E[X] = \sum_{i=1}^{N} i \cdot \frac{1}{N} = \frac{1}{N} \sum_{i=1}^{N} i = \frac{1}{N} \cdot \frac{N(N+1)}{2}$$

$$= \frac{N+1}{2}$$

$$E[X^2] = \sum_{i=1}^{N} i^2 \cdot \frac{1}{N} = \frac{1}{N} \frac{N(N+1)(2N+1)}{6}$$

$$= \frac{(N+1)(2N+1)}{6}$$

(13)

$$V[X] = \frac{(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2$$
$$= \frac{N^2-1}{12}$$

$$M_X(t) = E[e^{tX}] = \sum_{i=1}^{N} e^{ti} \cdot \frac{1}{N}$$
$$= \frac{1}{N} \sum_{i=1}^{N} e^{ti}$$

Def$^{\underline{n}}$: The <u>Bernoulli distribution</u>

x	$p(x)$
0	$1-p$
1	p

(14)

$$E[X] = 0(1-p) + 1 \cdot p = p$$
$$E[X^2] = 0^2(1-p) + 1^2 \cdot p = p$$
$$V[X] = p - p^2 = p(1-p) = pq \quad (q = 1-p)$$

$$M_X(t) = E[e^{tX}] = e^{t \cdot 0} \cdot q + e^{t \cdot 1} \cdot p$$
$$= pe^t + q$$

(15)

Binomial experiment

- Sequence of n independent trials
- each trial results in one of 2 possible outcomes (0,1)
- the probability of a "1" is the same on each trial (p)
- X counts the number of 1s

(16)

Binomial distribution

x	$p(x)$
0	q^n
1	$n p q^{n-1}$
2	$\binom{n}{2} p^2 q^{n-2}$
$\vdots$	
n	p^n

$$p(x) = \binom{n}{x} p^x q^{n-x}$$

We know from past examples that $\mu = np$

$$\sigma^2 = npq$$

$$M_X(t) = (pe^t + q)^n$$

Exam 2 question types:

1. Given the pdf of a random variable, find the pdf of a transformation of the random variable.
2. Given a continuous or discrete distribution, find the expected value and variance.
3. Given a continuous or discrete distribution, find the moment generating function.
4. Given a moment generating function, find the mean and variance.
5. Apply Chebyshev's Inequality to a given distribution.

Stat 561 HW 4

2.24 Compute $\mathrm{E}\,X$ and $\mathrm{Var}\,X$ for each of the following probability distributions.

(a) $f_X(x) = ax^{a-1}$, $0 < x < 1$, $a > 0$
(b) $f_X(x) = \frac{1}{n}$, $x = 1, 2, \ldots, n$, $n > 0$ an integer
(c) $f_X(x) = \frac{3}{2}(x-1)^2$, $0 < x < 2$

2.28 Let μ_n denote the nth central moment of a random variable X. Two quantities of interest, in addition to the mean and variance, are

$$\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}} \quad \text{and} \quad \alpha_4 = \frac{\mu_4}{\mu_2^2}.$$

The value α_3 is called the *skewness* and α_4 is called the *kurtosis*. The skewness measures the lack of symmetry in the pdf (see Exercise 2.26). The kurtosis, although harder to interpret, measures the peakedness or flatness of the pdf.

(a) Show that if a pdf is symmetric about a point a, then $\alpha_3 = 0$.
(b) Calculate α_3 for $f(x) = e^{-x}$, $x \geq 0$, a pdf that is *skewed to the right.*
(c) Calculate α_4 for each of the following pdfs and comment on the peakedness of each.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad -\infty < x < \infty$$

$$f(x) = \frac{1}{2}, \quad -1 < x < 1$$

$$f(x) = \frac{1}{2} e^{-|x|}, \quad -\infty < x < \infty$$

2.33 In each of the following cases verify the expression given for the moment generating function, and in each case use the mgf to calculate $\mathrm{E}\,X$ and $\mathrm{Var}\,X$.

(a) $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $M_X(t) = e^{\lambda(e^t - 1)}$, $x = 0, 1, \ldots;\ \lambda > 0$
(b) $P(X = x) = p(1-p)^x$, $M_X(t) = \frac{p}{1-(1-p)e^t}$, $x = 0, 1, \ldots;\ 0 < p < 1$
(c) $f_X(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}$, $M_X(t) = e^{\mu t + \sigma^2 t^2/2}$, $-\infty < x < \infty;\ -\infty < \mu < \infty$, $\sigma > 0$