Math 20, Fall 2017

Edgar Costa

Week 6

Dartmouth College

Exponential distribution

- $T = \text{Exp}(\lambda)$ (λ is any positive constant, depending on the experiment.)
- How long until something happens? (that occurs continuously and independently at a constant average rate)
  For example: time between occurrences of a Poisson processes (work it out!)
- $\Omega_T = [0, +\infty]$
- $f_T(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$
- $P(T \leq t) = 1 - e^{-\lambda t}$ if $t \geq 0$.
- $P(T > t + s | T \geq s) = P(T > t)$ (memoryless!)
- $E[T] = \frac{1}{\lambda}, \quad V[T] = \frac{1}{\lambda^2}$

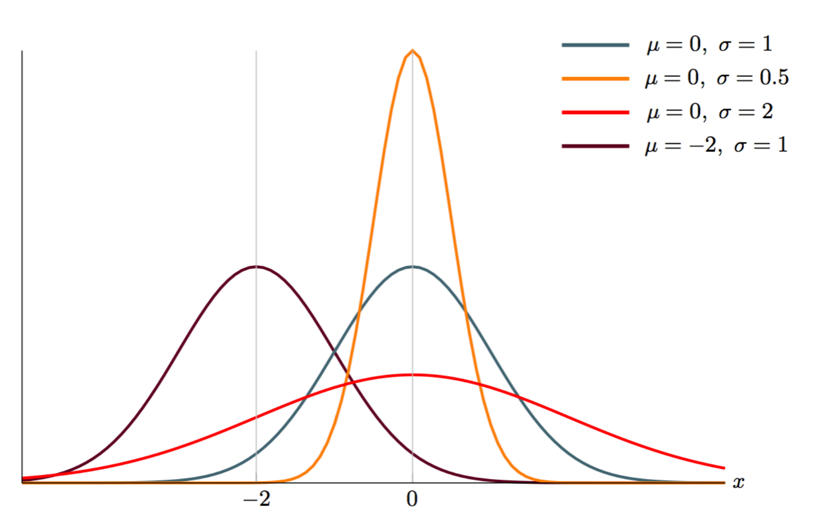
Normal distribution

- The normal density function of the normal distribution $N(\mu, \sigma)$ with parameters μ and σ is defined as follows:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}.$$

- The parameter μ represents the “center” of the density.
- The parameter σ is a measure of “spread” of the density, and thus it is assumed to be positive.

Normal distribution: In a picture



- We focus mostly on $\mu = 0$ and $\sigma = 1$
- We will call this particular normal density function the *standard normal density*, and we will denote it by $\phi(x)$:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

- There is no nice formula for $\int_a^b \phi(x)\,\mathrm{d}x$
- We instead use numerical tables for $\int_0^d \phi(x)\,\mathrm{d}x$
- Note that

$$\frac{N(\mu, \sigma) - \mu}{\sigma} \sim N(0, 1)$$

Exercise

On a test that determines whether an applicant receives a scholarship, the scores are distributed by a normal random variable with $\mu = 500$, $\sigma = 100$. It the top 5% of scores qualify for a scholarship, how high a score do you need to get it?

We seek a such that $P(X \geq a) = 0.05$. Then $P(X < a) = 0.95$.

For $Z = N(0, 1)$, we have $P(Z \leq 1.65) \approx 0.95$

So $\frac{a-\mu}{\sigma} = 1.65 \rightsquigarrow a = 665$

Exercise

Suppose that the height, in inches, of a 25-year old man is a normal random variable with parameters $\mu = 71$ and $\sigma^2 = 6.25$. What percentage of 25-year old men are over 6 feet 2 inches tall? What percentage of men over 6 feet tall are over 6 foot 5 inches?

Probability

- First, we defined probability in a intuitive way, as the frequency with which that outcome occurs in the long run.
- Later on, we defined probability mathematically as a value of a distribution function for the random variable representing the experiment.
- Now, with the law of large numbers, we will see that these two models are consistent.

Theorem

Let X be discrete random variable with expected value $\mu = E[X]$, and let $\epsilon > 0$ be any positive real number. Then

$$P(|X - \mu| \geq \epsilon) \leq \frac{V[X]}{\epsilon^2}$$

Proof:

$P(|X - \mu| \geq \epsilon) = \sum_{|x-\mu|\geq\epsilon} m_X(x)$

$$\epsilon^2 P(|X - \mu| \geq \epsilon) = \sum_{|x-\mu|\geq\epsilon} \epsilon^2 m_X(x) \leq \sum_{|x-\mu|\geq\epsilon} (x-\mu)^2 m_X(x) \leq \sum_{x\in\Omega} (x-\mu)^2 m_X(x)$$

Example

- Let X by any random variable with $E[X] = \mu$ and $V[X] = \sigma^2$
- If $\epsilon = k\sigma$, the Chebyshev Inequality states

$$P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{(k\sigma)^2} = \frac{1}{k^2}$$

- Thus, for any random variable, the probability of a deviation from the mean of more than k standard deviations is $\leq \frac{1}{k^2}$. .

Can we do better?

$$P(|X-\mu| \geq \epsilon) \leq \frac{V[X]}{\epsilon^2}$$

- Can we replace $\leq$ by strict inequality $<$?
- For a given ϵ is there a random variable X such that

$$P(|X-\mu| \geq \epsilon) = \frac{V[X]}{\epsilon^2}?$$

Chebyshev's Inequality

Theorem

Let X be discrete or continuous random variable with expected value $\mu = E[X]$, and let $\epsilon > 0$ be any positive real number. Then

$$P(|X - \mu| \geq \epsilon) \leq \frac{V[X]}{\epsilon^2}$$

Proof: For the continuous case replace in the appropriate manner $\sum$ by $\int$.

Exercises

$$P(|X - \mu| \geq \epsilon) \leq \frac{V[X]}{\epsilon^2}$$

- Let X be a random variable with $E[X] = 0$ and $V[X] = 1$. What integer value k will assure us that $P(|X| \geq k) \leq .01$?

Law of large numbers

Theorem

Let $X_1, X_2, \ldots, X_n$ be an independent trials process, with finite expected value $\mu = E[X_j]$ and finite variance $\sigma^2 = V[X_j]$. Let $S_n = X_1 + X_2 + \cdots + X_n$. Then for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \to 0$$

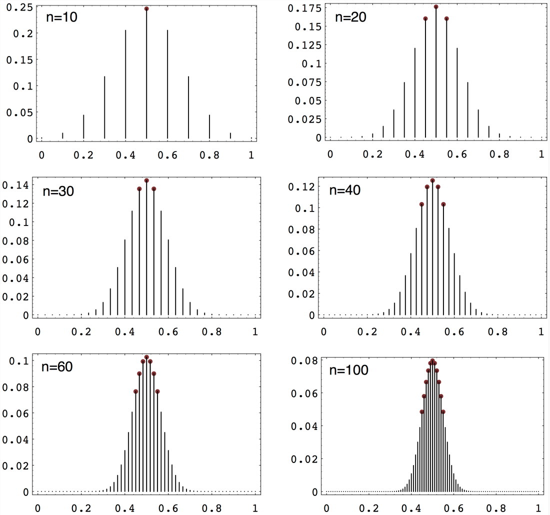
as $n \to +\infty$. Equivalently,

$$P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) \to 1$$

as $n \to +\infty$.

Prove it!

Binomial



Exercise

What do Chebyshev's Inequality and the Law of Large Numbers say about the probability of getting at least 75 heads when flipping a fair coin 100 times?
Hint: the binomial distribution is symmetric.