Chapter 5: Special Random Variables

ST2334 Probability and Statistics[1]
(Academic Year 2014/15, Semester 1)

Department of Statistics & Applied Probability (DSAP)
National University of Singapore (NUS)[1]

Outline

Introduction I

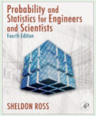
Learning Outcomes

Questions to Address: ♦ Understanding various common discrete/cont. r.v.'s ♦ Standardization of a normal r.v. ♦ How to compute normal probs & solve the "inverse" problem from Z-table ♦ How to read from Z-/χ^2-/t-/F-table

Concept & Terminology: ♦ binomial/hypergeometric/Poisson distribution ♦ exponential/normal/Gaussian/standard normal/chi-square distribution ♦ bell–/mound–shaped curve ♦ t-/F-distribution ♦ Z-/χ^2-/t-/F-table ♦ linear transformation/standardization of a normal r.v. ♦ cdf of the standard normal distribution ♦ percentile ♦ linear combination of independent normal r.v.'s

Introduction II

Mandatory Reading



✏ Section 5.1 – Section 5.8 (except Section 5.6.1, Section 5.7, & Section 5.8.1.1)

More Discrete r.v.'s

- Being the simplest experiment resulting in only 2 possible outcomes, different *Bernoulli trials appear very often in the real world*
- In the real world, there are *many more random phenomena* corresponding to *experiments defined systematically by or related to ≥ 2 Bernoulli trials*
  $\Rrightarrow$ *abundant applications of r.v.'s constructed from these experiments*
- Some of these r.v.'s are, namely,
  1. *binomial r.v.*
  2. *hypergeometric r.v.*
  3. *Poisson r.v.*
  4. *geometric r.v.* (not discussed here)
  5. *negative binomial r.v.* (not discussed here)

The Binomial Distribution I

- A Bernoulli trial associated with prob of success p (& prob of failure $q = 1 - p$) is repeatedly performed for n indept times ($n > 1$ is fixed)
- One usual interest is about X, the *total # of successes observed from the n trials*
- X is a *binomial r.v. with # of trials n & prob of success p* where p is the prob of observing a success from every Bernoulli trial

Definition

A r.v. X is said to have a *binomial distribution* with *parameters n & p*, write $X \sim Bin(n, p)$, if its pmf is defined by

$$p(x) = \binom{n}{x} p^x \, q^{n-x}, \qquad x = 0, 1, 2, \ldots, n$$

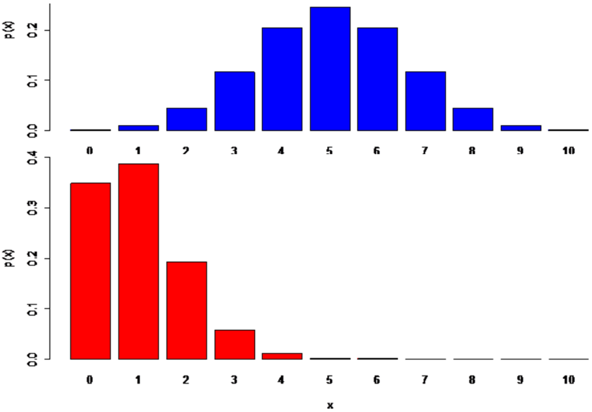
- $\sum_{x=0}^{n} p(x) = (p+q)^n = 1$ due to *binomial theorem*

The Binomial Distribution II

- The pmf $p(x) = P(X = x)$ for $x = 0, 1, 2, \ldots, n$ can be found by counting methods & standard prob computations:
  - Any outcome of the experiment is an ordered arrangement/a sequence of S & F
  - A particular sequence of x # of S's & $n - x$ # of F's occurs with prob $p^x q^{n-x}$ (following from multiplication law & indep between all the trial results)
  - $\binom{n}{x}$ # of ways to obtain combinations of x S's & $n - x$ F's
  - Adding up $p^x q^{n-x}$ for so many times $(= \binom{n}{x}$ times) gives $p(x)$ (following from addition law & disjointness of all these sequences)
- *Mean* $E(X) = np$; *Variance* $\mathrm{Var}(X) = npq$ (shown in Examples 13 & 20 (Ch.4), respectively)
- *Remark*: When we set $n = 1$ in the construction, we see that $Ber(p) \equiv Bin(1, p)$

The Binomial Distribution III

- The shape of pmf or “distribution” varies as a function of p: *symmetric* versus *skewed* as shown by pmf’s with $n = 10$, & $p = .5$ **(above)** or $p = .1$ **(below)**



Example 1: *Binomial Distribution*

Examine 3 light bulbs:

- X, total # of defective bulbs among 3 light bulbs, is of interest
- $n = 3$ identical Bernoulli trials resulting in 1 of 2 outcomes, {, } with prob of success () p: 3 bulbs to be examined ($\because$ each bulb is either or , & p is the true proportion of defective bulbs in the population of all bulbs)
- The 3 trials are indept ($\because$ defectiveness of a bulb does not affect one another)
- $X \sim Bin(3, p)$
- *Remark*: Here p is unknown to us but known to be a certain real # between 0 & 1

Example 2: Binomial Distribution

Win 1st prize in a draw of TOTO:

- X, the # of 1st prize winner, is of interest
- n identical Bernoulli trials resulting in 1 of 2 outcomes, $\{W, L\}$, with prob of success (W) p ($\because$ n is the total # of bets, & each bet results in either win or loss; $p = 1/\binom{45}{6} = \dfrac{1}{8,145,060}$ is the true prob of winning the 1st prize)
- All the n trials are indept ($\because$ winning of a bet does not affect the other bets)
- $X \sim Bin(n, \frac{1}{8,145,060})$
- *Remark*: Here n is unknown to us but known to be a fixed integer

Example 3: *Binomial Distribution*

Suppose that it is known that a manufacturer produces defective fuses subject to a prob of .05. In a lot of 100 produced fuses, what are the probs that

1. there are 2 defective fuses?
2. there are less than 5 defective fuses?

Solution: Let X be the # of defective fuses in a lot of size 100. Then, X is a binomial r.v. with parameters 100 & .05

$$P(\text{2 defective fuses}) = P(X = 2) = \binom{100}{2}(.05^2)(.95^{98}) = .0812$$

$$\begin{aligned} &P(< \text{5 defective fuses}) = P(X < 5) = P(X \leq 4) \\ &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ &= .436 \end{aligned}$$

The Poisson Distribution

Definition

A r.v. X is said to have a *Poisson distribution* with *parameter* $\lambda > 0$, write $X \sim Poi(\lambda)$, if its pmf is defined by

$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \qquad x = 0, 1, 2, \ldots$$

- λ is called the *average rate of occurrence of an "event"*
- $e \approx 2.71828$ denotes the *base of the natural logarithm system*
- $E(X) = Var(X) = \lambda$
- A *binomial r.v. with infinite # of trials*: Derived as the *limit of* $Bin(n, p)$
  - *Poisson approximation to binomial probs of* $Y \sim Bin(n, p)$:

$$P(Y = x) = \binom{n}{x} p^x q^{n-x} \approx p(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \qquad x = 0, 1, 2, \ldots, n$$

when $n \to \infty$ & $p \to 0$ s.t. $np = \lambda$ is moderate

Applications of Poisson Distribution

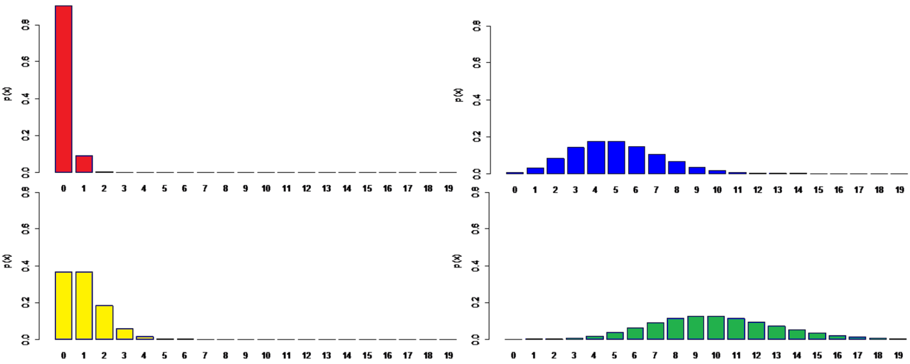
Poisson distribution is a *good model for the # of occurrence of a rare incidence* in a fixed period of time, in a given space, in a given volume, or on given physical objects:

- # of misprints on a page
- # of holes in an A4–paper after 10 years
- # of people in a community living to 100 years
- # of wrong telephone #'s that are dialed in a day
- # of people entering a store on a given day
- # of particles emitted by a radioactive source
- # of car accidents in a day
- # of people having a rare kind of disease

The Poisson pmf

$\lambda = .1$

$\lambda = 5$



$\lambda = 1$

$\lambda = 10$

Example 4: *Poisson Distribution*

Flaws (bad records) on a used video tape occur on the average of 1 flaw per 1, 200 feet. What are the probs that

1. there is no flaw on a tape of 4, 800 feet?
2. there are more than 2 flaws on a tape of 4, 800 feet?

Solution: Let X be the # of flaws on a tape of 4, 800 feet. Assume that $X \sim Poi(4 \times 1 = 4)$

1. $P(\text{no flaw}) = P(X = 0) = \dfrac{4^0}{0!}e^{-4} = e^{-4} = .018$
2. $P(> 2 \text{ flaws}) = P(X > 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$
   $= 1 - .018 - \dfrac{4^1}{1!}e^{-4} - \dfrac{4^2}{2!}e^{-4} = 1 - .238 = .762$

Example 5: *Poisson Approximation to Binomial Probabilities*

In a huge community, it is known that .7% of the population is color–blinded. What is the prob that at most 10 in a group of 1,000 people from the community are color–blinded?

Solution: Here, we assume that the # of color–blinded people in a group of 1,000 people from this community, $Y \sim Bin(1000, .007)$. The required prob is

$$P(Y \leq 10) = \sum_{k=0}^{10} \binom{1000}{k} (.007)^k (.993)^{1000-k} = .9022$$

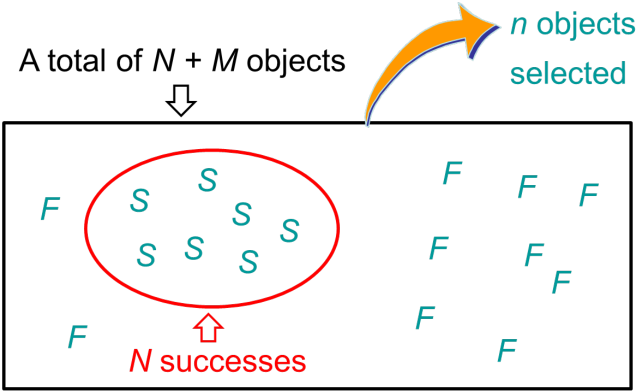
As $n = 1000$ is large & $p = .007 \approx 0$, *Y has approximately a Poisson distribution with parameter* $1,000 \times .007 = 7$. Hence,

$$P(Y \leq 10) \approx P(X \leq 10) = \sum_{k=0}^{10} \frac{7^k}{k!} e^{-7} = .9015$$

where $X \sim Poi(7)$

The Hypergeometric Distribution I

- From a population of 2 kinds of objects (S & F) with N # of S & M # of F, a total of $n < N + M$ objects are drawn *without replacement*



The Hypergeometric Distribution II

- One usual interest is about X, the *total # of S drawn*
- X is a *hypergeometric r.v. with parameters N, M, & n* where N is *# of successes*, M is *# of failures*, & n is *sample size*

Definition

A r.v. X is said to have a *hypergeometric distribution* with *parameters* N, M, n, write $X \sim Hyper(N, M, n)$, if its pmf is defined by

$$p(x) = \frac{\binom{N}{x}\binom{M}{n-x}}{\binom{N+M}{n}}$$

- x is an integer satisfying $\max(0, n - M) \leq x \leq \min(N, n)$
  - *Need to check*: Whether it is possible to obtain x S's together with $n - x$ F's at the same time

The Hypergeometric Distribution III

- $p(x)$ is derived easily due to the fact that the *order of the n selected objects does not matter*
  - total # of ways to sample n out of $N + M$ objects is $\binom{N+M}{n}$
  - total # of ways to sample x out of N S's together with $n - x$ out of M F's is given by $\binom{N}{x}\binom{M}{n-x}$
- *Mean* $E(X) = np$; *Variance* $\mathrm{Var}(X) = npq\left(1 - \frac{n-1}{N+M-1}\right)$ (with $p = N/(N+M)$ & $q = 1 - p$)
- *Hypergeometric versus binomial*:
  A *$Bin(n, p)$ r.v.* can be alternatively defined by the same mechanism/procedure as above with the sample size set as the # of trials n & $p = N/(N+M)$, *except that the objects are drawn with replacement*

Example 6: Hypergeometric Distribution

It is common that manufacturers perform quality control of their products by sampling a few products from a lot. When there are defective items more than a threshold value, then the lot will not be shipped. Suppose that in a lot of size 20, 4 of the products are defective. When there are > 2 defective items among 10 inspected, the lot will be rejected. What is the prob that this lot will be rejected?

Solution: Let X be the # of defective items observed in a sample of size 10. Then, X is a hypergeometric r.v. with parameters $N = 4$, $M = 16$ & $n = 10$, which takes on values $0, 1, 2, 3, 4$

$$
\begin{aligned}
P(\text{rejected}) = P(X > 2) &= P(X = 3) + P(X = 4) \\
&= \frac{\binom{4}{3}\binom{16}{7}}{\binom{20}{10}} + \frac{\binom{4}{4}\binom{16}{6}}{\binom{20}{10}} = .291
\end{aligned}
$$

The Exponential Distribution I

Definition

A r.v. X is called an *exponential r.v.* with *parameter* λ, write $X \sim Exp(\lambda)$, if, for $\lambda > 0$, its pdf is given by

$$f(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$

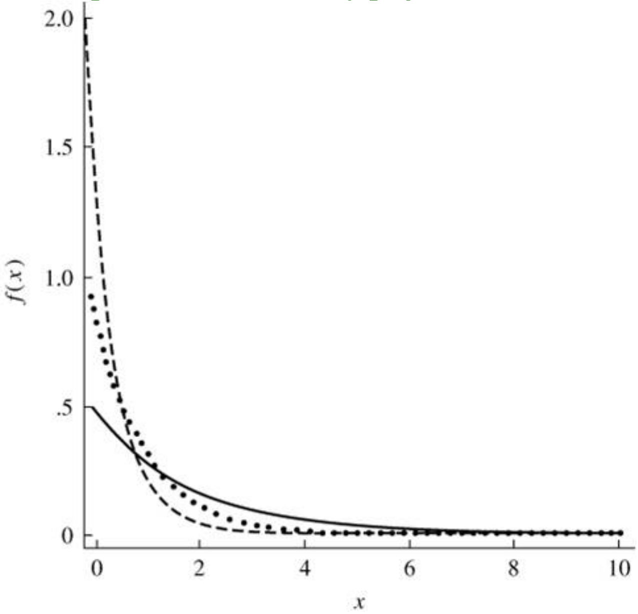
- 1 parameter $\lambda > 0$, like a Poisson r.v.
- *Family of exponential densities* indexed by different values of λ
- Common usage: Model lifetimes or waiting times until occurrence of some specific event (*e.g.,* earthquake, wrong phone call, & arrival of a new customer)
- The larger λ, the more rapidly the pdf drops off

The Exponential Density/pdf

$\lambda = .5$ **(solid)**

$\lambda = 1$ **(dotted)**

$\lambda = 2$ **(dashed)**



The Exponential Distribution II

- The *cdf of $X \sim Exp(\lambda)$* is given by

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Memorylessness Property

For $X \sim Exp(\lambda)$, & $s, t > 0$,

$$P(X > t + s | X > s) = P(X > t) = e^{-\lambda t}$$

is indept of s

- Think of X as being the lifetime of some creature: If the creature is *alive at age $t > 0$*, the distribution of the *remaining amount of time* that it survives is the *same as the original lifetime distribution* when it was at age 0
- The *only cont. r.v.* possessing this property

Example 7: Exponential Distribution

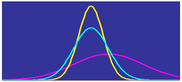
Suppose that the length of a phone call in minutes is an exponential r.v. with parameter $\lambda = .1$. Someone arrives immediately ahead of you at a public phone booth, find the probs that you will have to wait

1. more than 10 minutes,
2. between 10 to 20 minutes, &
3. more than 20 minutes after having already waited for 10 minutes

Solution: Let X be the duration of the person's call. Then, $X \sim Exp(.1)$

1. $P(X > 10) = 1 - F(10) = e^{-.1\times 10} = e^{-1} = .368$
2. $P(10 < X < 20) = F(20) - F(10) = (1 - e^{-2}) - (1 - e^{-1}) = .233$
3. $P(X > 20|X > 10) = P(X > 10) = .368$

Normal Random Variables I



Definition

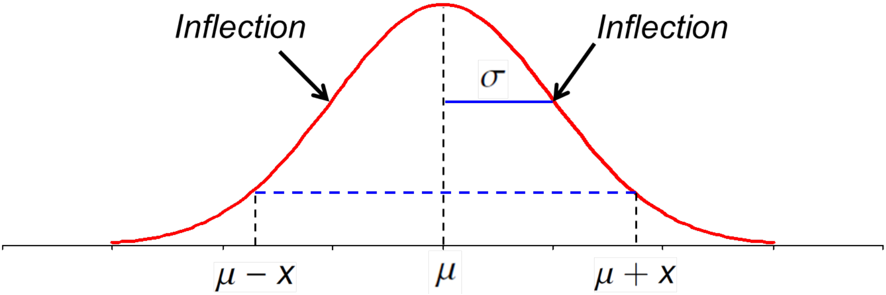
A r.v. X has a *normal or Gaussian distribution* with *parameters* μ *&* σ, write $X \sim N(\mu, \sigma^2)$, if, for $-\infty < \mu < \infty, \sigma > 0$, its pdf is given by

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

- Again, $e \approx 2.71828$ denotes the base of the natural logarithm system
- $\pi \approx 3.14159$ represents the familiar mathematical constant related to a circle
- *2 parameters* μ *&* σ^2: called the *mean* & the *variance* due to the fact that $E(X) = \mu$ & $Var(X) = \sigma^2$ (to be shown later in `Example 8`)

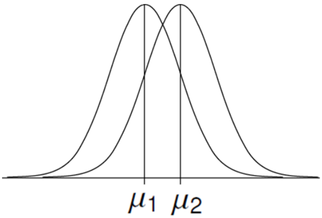
The Normal Density/pdf I

- $f(x)$ is a *bell–shaped/mound–shaped* curve
- $f(x)$ is *symmetric about μ* $\Rrightarrow f(\mu - x) = f(\mu + x)$ & $P(X < \mu) = P(X > \mu) = .5$
- *Points of inflection* at $x = \mu + \sigma$ & $x = \mu - \sigma$

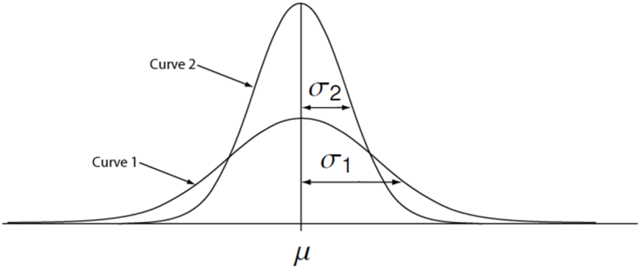


The Normal Density/pdf II

- μ: *center/mean* of the distribution; locates the *maximum/peak* of the curve; when $\mu_1 < \mu_2$



- σ: *shape/spread* of the distribution; the larger σ, the lower & the wider the curve; when $\sigma_2 < \sigma_1$



Normal Random Variables II

- *Play a central role in probability & statistics*
- The *most widely used models for diverse phenomena* as measurement errors in scientific experiments, reaction times in psychological experiments, measurements of intelligence & aptitude, scores on various tests, & so on
- Many *positive r.v's*, such as height, weight, & time, have distributions that are *well approximated by a normal distribution*
- *Central Limit Theorem (CLT)* to be introduced in `Ch.6` – *sum of indept r.v.'s is approximately normal* – justifies the use of the normal distribution in many applications

Linear Transformation of a Normal Random Variable

Suppose that $X \sim N(\mu, \sigma^2)$. Then, $Y = a + bX$, for fixed constants a & b, is a *normal r.v. with mean $a + b\mu$ & variance $b^2\sigma^2$*

- *A streamline proof for $b > 0$*: The cdf of Y is, for $y \in \mathbb{R}$,

$$F_Y(y) = P(Y \leq y) = P(a + bX \leq y) = P\left(X \leq \frac{y-a}{b}\right) = F_X\left(\frac{y-a}{b}\right)$$

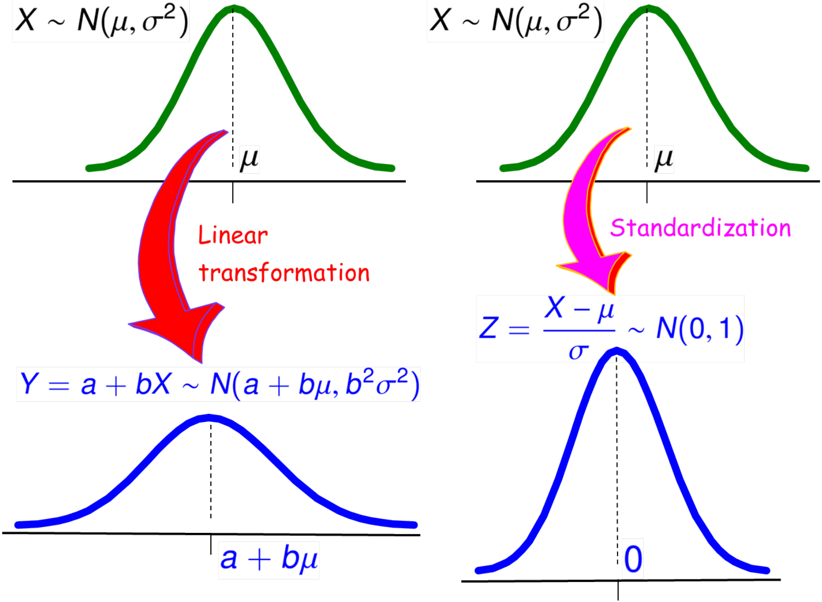
Thus, for $y \in \mathbb{R}$, the pdf of Y is

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{b}f_X\left(\frac{y-a}{b}\right) = \frac{1}{(b\sigma)\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-(a+b\mu)}{(b\sigma)}\right)^2\right]$$

$\therefore Y \sim N(a + b\mu, b^2\sigma^2)$

The *same result* can be obtained *for $b < 0$*

Linear Transformation of a Normal r.v. II



Standard Normal Random Variable

Standardization of a Normal r.v.

Suppose that $X \sim N(\mu, \sigma^2)$. Then, $Z = \dfrac{X - \mu}{\sigma}$ is a *standard normal r.v.* with mean 0 & variance 1

- Notice that $Z = -\frac{\mu}{\sigma} + \frac{1}{\sigma}X$ is a *linear transformation of X*, & apply the result at the previous page: Z is a normal r.v. with parameters

$$E(Z) = -\frac{\mu}{\sigma} + \frac{1}{\sigma}E(X) = -\frac{\mu}{\sigma} + \frac{1}{\sigma}(\mu) = 0$$

$$\mathrm{Var}(Z) = \left(\frac{1}{\sigma}\right)^2 \mathrm{Var}(X) = \left(\frac{1}{\sigma}\right)^2 \sigma^2 = 1$$

- The above *linear transformation* on the r.v. X defined by *subtracting the mean of X followed by dividing the result by the sd of X* is called *standardization* of X
- Usually reserve Z to denote a standard normal r.v.

Example 8: Mean & Variance of a Normal r.v.

For $Z \sim N(0,1)$, one can easily show that

$$E(Z) = \int_{-\infty}^{\infty} z f(z)\, dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-z^2/2}\, dz$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}\, d(z^2/2) = \frac{-1}{\sqrt{2\pi}} \left[e^{-z^2/2}\right]_{-\infty}^{\infty} = 0$$

&, using integration by parts,

$$\text{Var}(Z) = E(Z^2) - [E(Z)]^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2}\, dz - 0$$
$$= \frac{1}{\sqrt{2\pi}} \left([-ze^{-z^2/2}]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-z^2/2}\, dz \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}\, dz = 1$$

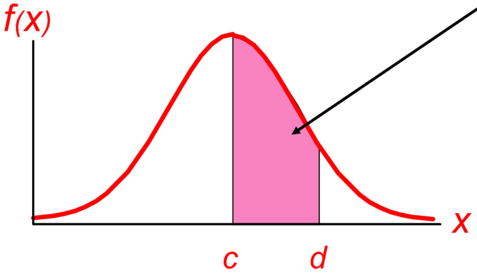
Hence, for $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$, we have

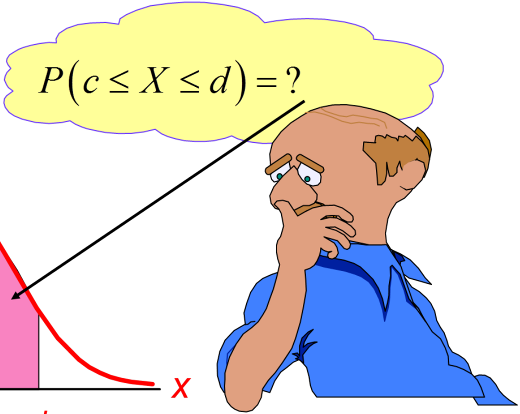
$$E(X) = \mu + \sigma E(Z) = \mu \qquad \& \qquad \text{Var}(X) = \text{Var}(\sigma Z) = \sigma^2 \text{Var}(Z) = \sigma^2$$

Computing Probabilities of Normal r.v.'s I

How do we compute probs for $X \sim N(\mu, \sigma^2)$?

Probability is the area under the density *f(x)*!





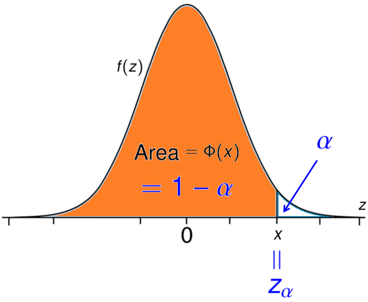
Computing Probabilities of Normal r.v.'s II

Definition

The *cdf of a $N(0,1)$ r.v.*, Z, is defined by, for $-\infty < x < \infty$,

$$\Phi(x) = P(Z \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \tag{1}$$

- *Area under the standard normal density* $f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ *between $-\infty$ & x*
  $\Rightarrow \Phi(-\infty) = 0, \Phi(0) = .5, \Phi(\infty) = 1$
- *No closed-form expression* for the integral in (1)
- *Z-table* (shown at `Page 36`; click DOWNLOAD to download):
  1. tabulate *values of $\Phi(x)$* for other values for $x \ge 0$
  2. for $-x > 0$, $\Phi(-x) = 1 - \Phi(x)$



Computing Probabilities of Normal r.v.'s III

Definition

For any $0 < \alpha < 1$, the *value* z_α such that $\Phi(z_\alpha) = P(Z < z_\alpha) = 1 - \alpha$ is called the $100(1-\alpha)$ *percentile* of the standard normal distribution

cdf of a $N(\mu, \sigma^2)$ r.v. in terms of $\Phi(\cdot)$

The *cdf of* $X \sim N(\mu, \sigma^2)$ is given by

$$F(x) = P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

for $-\infty < x < \infty$

Computing Probabilities of a $N(\mu, \sigma^2)$ r.v.

For $X \sim N(\mu, \sigma^2)$,

$$P(c \leq X \leq d) = P\left(\frac{c-\mu}{\sigma} \leq Z \leq \frac{d-\mu}{\sigma}\right) = \Phi\left(\frac{d-\mu}{\sigma}\right) - \Phi\left(\frac{c-\mu}{\sigma}\right)$$

for $-\infty < c \leq d < \infty$

Z-Table

TABLE A.1 *Standard Normal Distribution Function:* $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\,dy$

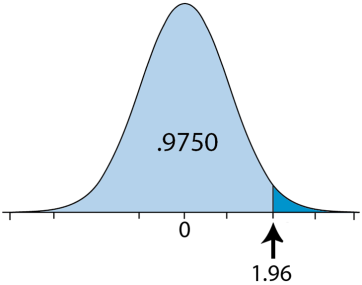
x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633

Example 9: *Lower–tailed Probabilities of* $Z \sim N(0,1)$

Lower–tailed probs:

1. $P(Z \le 1.96) = \Phi(1.96) = .9750$ ($\Rightarrow z_{.025} = 1.96$)
2. $P(Z \le .34) = \Phi(.34) = .6331$ ($\Rightarrow z_{.3669} = .34$)
3. $P(Z \le -1.96) = \Phi(-1.96) = 1 - \Phi(1.96) = .0250$ ($\Rightarrow z_{.975} = -1.96$)

	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750



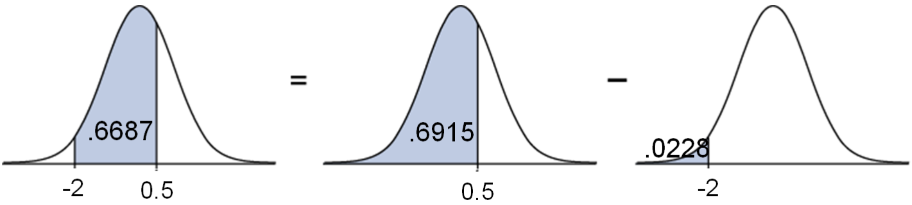
Example 10: *More Probabilities of* $Z \sim N(0,1)$

Upper–tailed probs:

1. $P(Z > 1.96) = 1 - P(Z \leq 1.96) = 1 - \Phi(1.96) = .0250$
2. $P(Z \geq -1.96) = 1 - P(Z < -1.96) = 1 - P(Z \leq -1.96)$
   $= 1 - \Phi(-1.96) = 1 - [1 - \Phi(1.96)] = .9750$

Probs between 2 points:

$$P(-2 \leq Z \leq .5) \quad = \quad \Phi(.5) \quad - \quad \Phi(-2)$$



where $\Phi(-2) = 1 - \Phi(2) = 1 - .9772 = .0228$

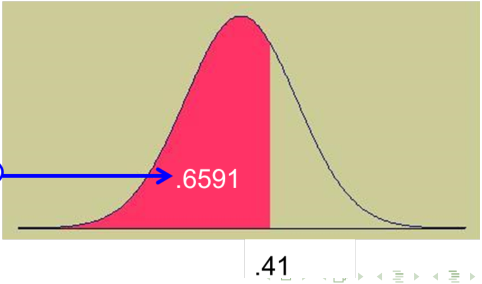
Example 11: *Computing Probabilities of a* $N(\mu, \sigma^2)$ *r.v.*

Let X be gestational length in weeks. We know from prior research: $X \sim N(39.18, 2^2)$. What is the proportion of gestations being less than 40 weeks? What is the prob that a pregnant woman will deliver less than 40 weeks?

Solution: For both questions, the answer is given by

$$P(X < 40) = P\left(Z < \frac{40 - 39.18}{2}\right) = P(Z < .41) = .6591$$

X	.00	.01
.0	.5000	.5040
.1	.5398	.5438
.2	.5793	.5832
.3	.6179	.6217
.4	.6554	.6591
.5	.6915	.6950
.6	.7257	.7291



The "Inverse" Problem

Definition

When $X \sim N(\mu, \sigma^2)$, we may be interested in finding the value of d s.t. $P(X \le d) = p$ for a given $0 < p < 1$. We refer to this problem as the *"inverse" problem* of computing probs of a normal r.v.

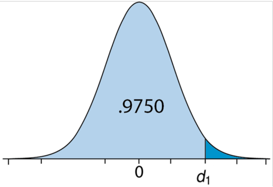
- **Idea**: When a normally distributed r.v. X is of interest, what is *the value of d* in the following claim *for any given* $0 < p < 1$

Proportion of time that a realization from X is $\le d$ is p

- *Locate p in the pool of #'s in the Z-table*, & then compute d based on the x value associated with p
  - In practice, the exact value *p may not be found* in the Z-table, *i.e.,* one can only find 2 #'s closest to p, a & b, s.t. $a < p < b$ $\Rrightarrow$ In this module, we will take *the most nearby #* from a & b as a *proxy of p* (See `Example 13`)

Example 12: "Inverse" Problem for $Z \sim N(0,1)$

Refer to Example 9, in which we obtain $\Phi(1.96) = .9750$, & $\Phi(.34) = .6331$. We can find the values of d_i in



❶ $P(Z \leq d_1) = \Phi(d_1) = .9750$

❷ $P(Z \leq d_2) = \Phi(d_2) = .6331$

❸ $P(Z \leq d_3) = \Phi(d_3) = .0250$

as $d_1 = 1.96$, $d_2 = .34$ & $d_3 = -1.96$

- For ❶ & ❷, one can just locate .9750 & .6331 from the Z-table, & then read the values of d_1 & d_2
- For ❸, it is *impossible* to locate .0250 in the Z-table as $d_3 < 0$. One has to work out the following equivalent equations:

$$\Phi(d_3) = 1 - \Phi(-d_3) = .0250 \quad \Leftrightarrow \quad \Phi(-d_3) = .9750$$

to conclude that $-d_3 = 1.96$ (*i.e.*, $d_3 = -1.96$)

Example 13: "Inverse" Problem for a $N(\mu, \sigma^2)$ r.v. I

Let X be gestational length in weeks. We know from prior research: $X \sim N(39, 2^2)$. What is the gestational length ℓ s.t. 40% of all gestation lengths are shorter?

Solution: When all the gestational lengths are known & collected, 40% of all measurements would be less than ℓ, *i.e.*,

$$P(X < \ell) = .4 \quad \Rrightarrow \quad P\left(Z < \frac{\ell - 39}{2}\right) = .4 \quad \Rrightarrow \quad \boxed{\Phi(d) = .4}$$

where $d = (\ell - 39)/2$ $\therefore$ It *suffices to find d for $p = .4$*

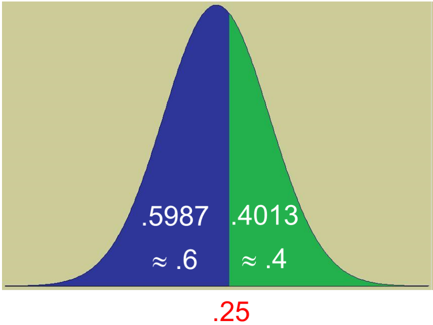
- Noticing that $p < .5$ implies that $d < 0$
- Re-write the above equation as $1 - \Phi(-d) = .4$ or $\underline{\Phi(-d) = .6}$ & *locate $p = .6$* from the Z-table
- As $\underline{\Phi(.25) = .5987 \approx .6}$ (compared with $\Phi(.26) = .6026$, .5987 is *closer to .6* than .6026), we equate $-d$ with .25 to get $d = -.25$, & thus,

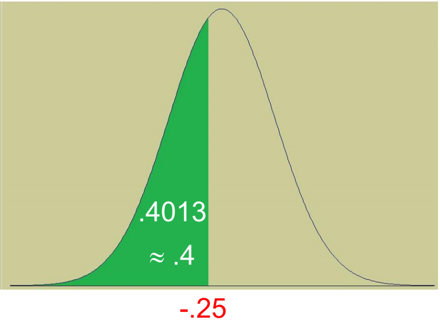
$$\ell = 39 + 2d = 39 + 2(-.25) = 38.5$$

Example 13: "Inverse" Problem for a $N(\mu, \sigma^2)$ r.v. II

$P(Z \le .25) = .5987 \approx .6$

$P(Z \le -.25) = P(Z \ge .25)$
$= 1 - P(Z < .25)$
$= 1 - .5987 = .4013 \approx .4$





A Summary of Properties of Some r.v.'s

	Probability mass function, $p(x)$	Moment generating function, $\phi(t)$	Mean	Variance
Binomial with parameters n, p; $0 \leq p \leq 1$	$\binom{n}{x} p^x(1-p)^{n-x}$ $x = 0, 1, \ldots, n$	$(pe^t + 1 - p)^n$	np	$np(1-p)$
Poisson with parameter $\lambda > 0$	$e^{-\lambda}\frac{\lambda^x}{x!}$ $x = 0, 1, 2, \ldots$	$\exp\{\lambda(e^t - 1)\}$	λ	λ

	Probability density function, $f(x)$	Moment generating function, $\phi(t)$	Mean	Variance
Uniform over (a, b)	$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential with parameter $\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\frac{\lambda}{\lambda - t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normal with parameters (μ, σ^2)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$ $-\infty < x < \infty$	$\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$	μ	σ^2

Sums & Linear Combinations of Indept Normal r.v.'s I

- The normal distribution possesses the *reproductive* property that the *sum of indept normal r.v.'s is also a normal r.v.*

Linear Combinations of Independent Normal r.v.'s

If X_i, $i = 1, \ldots, n$ are indept normally distributed r.v.'s with respective parameters μ_i & σ_i^2, $i = 1, \ldots, n$, then, for real constants $a, b_1, \ldots, b_n \in \mathbb{R}$,

$$Y = a + \sum_{i=1}^{n} b_i X_i \sim N\left(a + \sum_{i=1}^{n} b_i \mu_i, \sum_{i=1}^{n} b_i^2 \sigma_i^2\right)$$

- **Recall** (at Page 29) that linear transformation of a normal r.v. is also a normal r.v. $\Rrightarrow$ Each summand $b_i X_i$ at the LHS is also a normal r.v.

Sums & Linear Combinations of Indept Normal r.v.'s II

- A proof using mgf's:
  1. For $X \sim N(\mu, \sigma^2)$, its *mgf* $M(t)$ equals

$$
\begin{aligned}
E(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-[x^2+\mu^2-2(\mu+\sigma^2 t)x]/(2\sigma^2)}\, dx \\
&= e^{-\mu^2/(2\sigma^2)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-[x^2-2(\mu+\sigma^2 t)x]/(2\sigma^2)}\, dx \\
&= e^{-\mu^2/(2\sigma^2)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}\{[x-(\mu+\sigma^2 t)]^2-(\mu+\sigma^2 t)^2\}}\, dx \\
&= e^{\frac{1}{2\sigma^2}[-\mu^2+(\mu+\sigma^2 t)^2]} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}[x-(\mu+\sigma^2 t)]^2}\, dx = e^{\mu t} e^{\sigma^2 t^2/2}
\end{aligned}
$$

2. the mgf of $Y = a + \sum_{i=1}^{n} b_i X_i$, is given by

$$\begin{aligned}
M_Y(t) &= e^{at} M_{\sum_{i=1}^{n} b_i X_i}(t) \\
&= e^{at} M_{b_1 X_1}(t) \cdot M_{b_2 X_2}(t) \cdots M_{b_n X_n}(t) \qquad \text{[By (7) of Ch.4]} \\
&= e^{at} M_{X_1}(b_1 t) \cdot M_{X_2}(b_2 t) \cdots M_{X_n}(b_n t) \qquad \text{[By (6) of Ch.4]} \\
&= e^{at} e^{\mu_1 (b_1 t)} e^{\sigma_1^2 (b_1 t)^2/2} \cdot e^{\mu_2 (b_2 t)} e^{\sigma_2^2 (b_2 t)^2/2} \cdots e^{\mu_n (b_n t)} e^{\sigma_n^2 (b_n t)^2/2} \\
&= e^{(a + b_1 \mu_1 + \cdots + b_n \mu_n) t} e^{(b_1^2 \sigma_1^2 + \cdots + b_n^2 \sigma_n^2) t^2/2}
\end{aligned}$$

$\therefore$ Comparing it with the mgf of a $N(\mu, \sigma^2)$ r.v. at the previous page yields the result

Example 14: Sum of 2 Indept Normal r.v.'s I

The profit of a company is determined by the net profit in 2 of its sales outlet. Suppose that the net monthly profits (in 1,000 dollars), X_1 & X_2, in the 2 outlets are normally distributed with parameters 10.4 & 6.24, & 12.6 & 3.78, respectively. Compute the probs that

1. the company earns more than $25,000 in a month
2. the net profit in the 1st outlet is higher in this month

Solution:

1. $X_1 \sim N(10.4, 6.24)$ & $X_2 \sim N(12.6, 3.78)$
   $\Rightarrow X_1 + X_2 \sim N(10.4 + 12.6 = 23, 6.24 + 3.78 = 10.02)$. Hence,

$$P(X_1 + X_2 > 25) = P\left(\frac{X_1 + X_2 - 23}{\sqrt{10.02}} > \frac{25 - 23}{\sqrt{10.02}}\right)$$
$$= P\left(Z > .6318\right) \approx 1 - \Phi(.63) = .2643$$

Example 14: Sum of 2 Indept Normal r.v.'s II

2. Note that $X_1 - X_2 \sim N(10.4 - 12.6 = -2.2, 6.24 + 3.78 = 10.02)$. So, we have

$$P(X_1 - X_2 > 0) = P\left(\frac{X_1 - X_2 + 2.2}{\sqrt{10.02}} > \frac{0 + 2.2}{\sqrt{10.02}}\right)$$
$$= P\left(Z > .695\right) \approx 1 - \Phi(.70) = .2420$$

Regarding the mean of $X_1 - X_2$, $E(X_1 - X_2) = -2.2$, the profit of the 2nd outlet is expected to be 2,200 dollars per month more than that of the 1st outlet

Distributions Arising From The Normal

- **Recall**: The normal distribution is the most commonly used cont. prob distribution in probability & statistics
  - a lot of random phenomena can be approximately modeled by a normal distribution
- A few *specific functions of the normal distribution* also *occur in various statistical problems*
- These r.v.'s derived from the normal distribution are called
  1. *chi-square distribution*
  2. *t-distribution*
  3. *F-distribution*
  4. Cauchy distribution (not discussed here)

The Chi-Square Distribution I

Definition

If $Z \sim N(0,1)$, the *nonegative* r.v.

$$U = g(Z) = Z^2$$

is said to have a *chi-square distribution with 1 degree of freedom (df)*. We write $U \sim \chi_1^2$

- Its usual occurrence: $Y \sim N(\mu, \sigma^2) \Leftrightarrow (Y - \mu)/\sigma \sim N(0,1)$ $\Rightarrow [(Y-\mu)/\sigma]^2 \sim \chi_1^2$

Definition

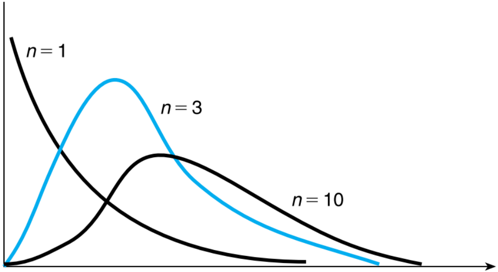
If $U_1, \ldots, U_n \overset{\text{iid}}{\sim} \chi_1^2$, the *nonegative* r.v.

$$X = g(U_1, \ldots, U_n) = U_1 + \cdots + U_n$$

is said to have a *chi-square distribution with n degrees of freedom*. We write $X \sim \chi_n^2$

The Chi-Square Distribution II

- The *df* parameter of a χ^2_n r.v. must be a *positive integer*
- The *pdf* of a χ^2_n r.v. is plotted as



- *Mean* $E(X) = n$; *Variance* $\mathrm{Var}(X) = 2n$
- Sum of indept chi-square r.v.'s:

$$\chi^2_m + \chi^2_n \equiv \chi^2_{m+n}$$

The t-Distribution I

Definition

If $Z \sim N(0,1)$ & $X \sim \chi^2_n$ are indept r.v.'s, the r.v.

$$T = g(X, U) = \frac{Z}{\sqrt{X/n}}$$

is said to have the *t-distribution with n degrees of freedom*, where n is a positive integer. We write $T \sim t_n$

- *Mean* $E(T) = 0, \ n > 1$; *Variance* $\mathrm{Var}(T) = \dfrac{n}{n-2}, \ n > 2$
- The pdf (plotted next page) of a t-distribution:
  - *symmetric about zero* as $f(t) = f(-t)$
  - has a *heavier tail* compared with the standard normal density

The t-Distribution II

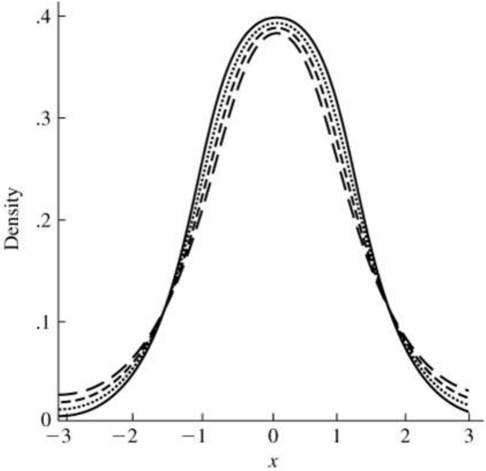
- As the *df approaches* ∞, the *t-distribution tends to the standard normal distribution Z*; in fact, for >20 or >30 df, the 2 distributions are very close

long dashes: $df = 5$

short dashes: $df = 10$

dots: $df = 30$

solid line: $N(0, 1)$



The F-distribution

Definition

If $U \sim \chi^2_n$ & $V \sim \chi^2_m$ are indept r.v.'s, the *nonegative* r.v.

$$W = g(U, V) = \frac{(U/n)}{(V/m)}$$

is said to have the *F-distribution with n & m degrees of freedom*, where m & n are both positive integers. We write $W \sim F_{n,m}$

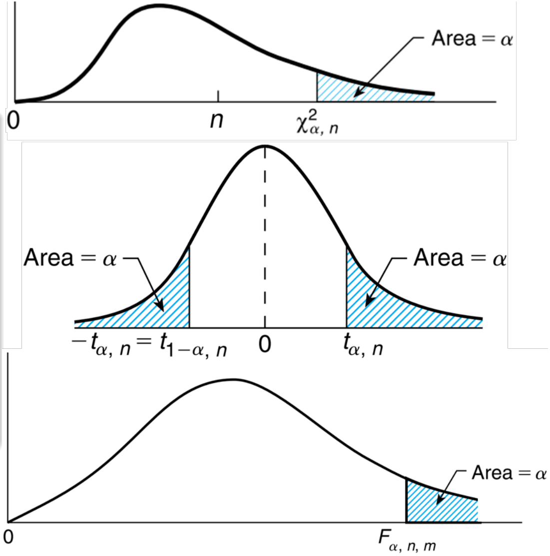
- *Mean* $E(W) = \dfrac{m}{m-2}, \; m > 2;$

  Variance $\mathrm{Var}(W) = 2\left(\dfrac{m}{m-2}\right)^2 \dfrac{n+m-2}{n(m-4)}, \; m > 4$
- $F_{1,m} \equiv t^2_m$

Percentiles of χ_n^2, t_n, And $F_{n,m}$ r.v.'s I

Definition

For $\alpha \in (0, 1)$, let $\chi^2_{\alpha,n}$ (*resp.*, $t_{\alpha,n}$ & $F_{\alpha,n,m}$) be the 100$(1-\alpha)$ *percentile* of the χ^2_n (*resp.*, t_n & $F_{n,m}$) distribution, which is the value of the point beyond which the χ^2_n (*resp.*, t_n & $F_{n,m}$) r.v. has prob α



Percentiles of χ^2_n, t_n, And $F_{n,m}$ r.v.'s II

- These percentiles can be *obtained from* the next 3 prob tables, *χ^2-table, t-table, & F-table*

Some Properties For Probabilities & Percentiles of The χ^2_n, t_n, & $F_{n,m}$ Distributions

- $P(\chi^2_n > \chi^2_{\alpha,n}) = \alpha$
- $P(t_n > t_{\alpha,n}) = P(t_n < -t_{\alpha,n}) = P(t_n < t_{1-\alpha,n}) = \alpha$
- $P\left(F_{n,m} > F_{\alpha,n,m}\right) = P\left(F_{n,m} > \frac{1}{F_{1-\alpha,m,n}}\right) = \alpha$

χ^2-*Table*

TABLE A.2 *Values of* $x^2_{\alpha,n}$

n	$\alpha = .995$	$\alpha = .99$	$\alpha = .975$	$\alpha = .95$	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
1	.0000393	.000157	.000982	.00393	3.841	5.024	6.635	7.879
2	.0100	.0201	.0506	.103	5.991	7.378	9.210	10.597
3	.0717	.115	.216	.352	7.815	9.348	11.345	12.838
4	.207	.297	.484	.711	9.488	11.143	13.277	14.860
5	.412	.554	.831	1.145	11.070	12.832	13.086	16.750
6	.676	.872	1.237	1.635	12.592	14.449	16.812	18.548
7	.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718

t-Table

TABLE A.3 *Values of* $t_{\alpha,n}$

n	$\alpha = .10$	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.474	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898

F-Table

TABLE A.4 *Values of* $F_{.05,n,m}$

m = Degrees of Freedom for Denominator	n = Degrees of Freedom for Numerator 1	2	3	4	5
1	161	200	216	225	230
2	18.50	19.00	19.20	19.20	19.30
3	10.10	9.55	9.28	9.12	9.01
4	7.71	6.94	6.59	6.39	6.26
5	6.61	5.79	5.41	5.19	5.05
6	5.99	5.14	4.76	4.53	4.39
7	5.59	4.74	4.35	4.12	3.97
8	5.32	4.46	4.07	3.84	3.69
9	5.12	4.26	3.86	3.63	3.48
10	4.96	4.10	3.71	3.48	3.33
11	4.84	3.98	3.59	3.36	3.20
12	4.75	3.89	3.49	3.26	3.11
13	4.67	3.81	3.41	3.18	3.03
14	4.60	3.74	3.34	3.11	2.96
15	4.54	3.68	3.29	3.06	2.90

Example 15: Probabilities of χ^2_n, t_n, And $F_{n,m}$ r.v.'s I

1. χ^2_n prob statements & percentiles from the χ^2-table (at Page 58):

$$
\begin{aligned}
P(\chi^2_1 < 3.841) = .95 &\Leftrightarrow \chi^2_{.05,1} = 3.841\ (= 1.96^2 = z^2_{.025}) \\
P(\chi^2_5 < 1.145) = .05 &\Leftrightarrow \chi^2_{.95,5} = 1.145 \\
P(\chi^2_5 < 12.832) = .975 &\Leftrightarrow \chi^2_{.025,5} = 12.832 \\
P(\chi^2_5 < 16.750) = .995 &\Leftrightarrow \chi^2_{.005,5} = 16.750
\end{aligned}
$$

   - $P(3.053 < \chi^2_{11} < 24.725) = P(\chi^2_{11} < 24.725) - P(\chi^2_{11} \leq 3.053)$
     $= (1 - .01) - (1 - .99) = .98$

2. t_n prob statements & percentiles from the t-table (at Page 59):

$$
\begin{aligned}
P(t_1 < 3.078) = .90 &\Leftrightarrow t_{.10,1} = 3.078 \\
P(t_9 < 2.262) = .975 &\Leftrightarrow t_{.025,9} = 2.262 \\
P(t_9 < -2.262) = P(t_9 > 2.262) &\Leftrightarrow t_{.975,9} = -2.262 \\
= 1 - P(t_9 < 2.262) = .025 &
\end{aligned}
$$

Example 15: Probabilities of χ^2_n, t_n, And $F_{n,m}$ r.v.'s II

- $P(1.383 < t_9 < 2.262) = P(t_9 < 2.262) - P(t_9 \le 1.383)$
  $= .975 - (1 - .10) = .075$
- $P(-1.746 < t_{16} < 1.746) = 1 - 2P(t_{16} \ge 1.746) = 1 - 2(.05) = .9$

3. $F_{n,m}$ prob statements & percentiles from the F-table (at Page 60):

$$\begin{aligned}
P(F_{1,1} > 161) = .05 &\quad\Leftrightarrow\quad F_{.05,1,1} = 161 \; (= 12.706 = t^2_{.025,1}) \\
P(F_{1,15} > 4.54) = .05 &\quad\Leftrightarrow\quad F_{.05,1,15} = 4.54 \\
P(F_{3,5} > 5.41) = .05 &\quad\Leftrightarrow\quad F_{.05,3,5} = 5.41 \\
P\left(F_{5,3} > \frac{1}{5.41}\right) = .95 &\quad\Leftrightarrow\quad F_{.95,5,3} = \frac{1}{F_{.05,3,5}} = \frac{1}{5.41} = .1848
\end{aligned}$$