Ch. 5 – Joint Probability Distributions

Joint Probability Mass Function

Let X and Y be discrete random variables. Their **joint probability mass function** (joint pmf) is

$$f(x, y) = P(X = x \cap Y = y)$$

Example: A insurance agency serves many customers with both an automobile policy and a homeowner's policy. The possible amounts for a deductible are \$100 and \$250 for an automobile policy, and \$0, \$100, and \$200 for a homeowner's policy. If a customer is selected at random, and X is their deductible on the auto policy and Y is their deductible on the homeowner's policy, then the joint pmf of X and Y may be described by a table such as the following:

$f(x, y)$	$y = 0$	$y = 100$	$y = 200$
$x = 100$	.20	.10	.20
$x = 250$	.05	.15	.30

Example

Suppose we roll two fair 6-sided dice, and let X be the smaller of the two rolls, and let Y be the larger. Find the joint probability mass function $f(x,y)$ of X and Y.

$f(x,y)$	$y=1$	$y=2$	$y=3$	$y=4$	$y=5$	$y=6$
$x=1$	$1/36$	$2/36$	$2/36$	$2/36$	$2/36$	$2/36$
$x=2$	0	$1/36$	$2/36$	$2/36$	$2/36$	$2/36$
$x=3$	0	0	$1/36$	$2/36$	$2/36$	$2/36$
$x=4$	0	0	0	$1/36$	$2/36$	$2/36$
$x=5$	0	0	0	0	$1/36$	$2/36$
$x=6$	0	0	0	0	0	$1/36$

For instance, $f(2,3) = 2/36$, but $f(3,2) = 0$.

Marginal Probability Mass Function

Let X and Y be discrete random variables with joint probability mass function $f(x, y)$. The **marginal probability mass function** of X is defined as

$$f_X(x) = \sum_y f(x, y)$$

where the sum is taken over all possible values of Y. Similarly, the marginal probability mass function of Y is

$$f_Y(y) = \sum_x f(x, y)$$

where the sum is taken over all possible values of X.

Note: f_X is simply the probability mass function of X considered as a random variable on its own: $f_X(x) = P(X = x)$.

Example

Recall the joint pmf for the auto insurance deductible X and homeowner's insurance deductible Y from earlier:

$f(x,y)$	$y=0$	$y=100$	$y=200$
$x=100$	.20	.10	.20
$x=250$	.05	.15	.30

The marginal pmf for the auto insurance deductible X is given by

$$f_X(100) = .50$$
$$f_X(250) = .50$$

The marginal pmf for the homeowner's deductible Y is given by

$$f_Y(0) = .25$$
$$f_Y(100) = .25$$
$$f_Y(200) = .50$$

Independence of Random Variables

Discrete random variables X and Y are **independent** if their joint probability mass function $f(x, y)$ is the product of the marginal probability mass functions $f_X(x)$ and $f_Y(y)$:

$$f(x, y) = f_X(x) f_Y(y)$$

In the insurance example,

$f(x, y)$	$y = 0$	$y = 100$	$y = 200$
$x = 100$	.20	.10	.20
$x = 250$	.05	.15	.30

the random variables X and Y are dependent, because, e.g.,

$$f(100, 0) = .20 \neq (.50)(.25) = f_X(100) f_Y(0)$$

Example

If X and Y are independent geometric random variables with parameter p and q respectively, find the joint pmf of X and Y.

Solution: The marginal pmf of X is

$$f_X(x) = p(1-p)^x$$

while the marginal pmf of Y is

$$f_Y(y) = q(1-q)^y$$

Since X and Y are independent, their joint pmf is the product of their marginal pmfs:

$$f(x,y) = f_X(x)f_Y(y) = p(1-p)^x q(1-q)^y$$

Sum of Two Discrete Random Variables

Given random variables X and Y with joint pmf $f(x, y)$, the sum $X + Y$ has pmf

$$f_{X+Y}(t) = \sum f(x, t - x)$$

where the sum is taken over all possible values x of X.

In particular, if X and Y are independent, then

$$f_{X+Y}(t) = \sum f_X(x) f_Y(t - x)$$

Example

Suppose X and Y are independent geometric random variables with parameters $p = .3$ and $q = .4$ respectively. What is the probability that $X + Y = 2$?

Solution: The marginal pmfs are $f_X(x) = p(1-p)^x$ and $f_Y(y) = q(1-q)^y$. Therefore,

$$
\begin{aligned}
&P(X+Y=2)\\
&= \sum_x f_X(x) f_Y(2-x)\\
&= f_X(0)f_Y(2) + f_X(1)f_Y(1) + f_X(2)f_Y(0)\\
&= p(1-p)^0 q(1-q)^2 + p(1-p)^1 q(1-q)^1 + p(1-p)^2 q(1-q)^0\\
&= (.3)(.4)(.6)^2 + (.3)(.7)(.4)(.6) + (.3)(.7)^2(.4) = .1524
\end{aligned}
$$

Expected Value

Given random variables X and Y with joint pmf $f(x, y)$, and given a function $h(x, y)$, the **expected value** of $h(X, Y)$ is

$$E[h(X, Y)] = \sum_x \sum_y h(x, y) f(x, y)$$

where the sums are over the possible values for X and Y.

Example: If X and Y are the auto and homeowner's insurance deductible as before, find the expected value of $|X - Y|$:

$$\begin{aligned} E(|X - Y|) &= \sum_x \sum_y |x - y| f(x, y) \\ &= |100 - 0|(.20) + |100 - 100|(.10) + |100 - 200|(.20) \\ &\quad + |250 - 0|(.05) + |250 - 100|(.15) + |250 - 200|(.30) \\ &= 90 \end{aligned}$$

Several Random Variables

If $X_1, \ldots, X_n$ are random variables, their joint pmf is defined as

$$f(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n)$$

The marginal pmf of X_1 is defined as

$$f_{X_1}(x) = \sum_{x_2} \sum_{x_3} \cdots \sum_{x_n} P(X_1 = x, X_2 = x_2, \ldots, X_n = x_n)$$

The random variables $X_1, \ldots, X_n$ are independent if

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

Problem

A manufacturer receives components of a certain type from 3 plants: 50% from plant A, 30% from plant B, 20% from plant C. If you randomly select 9 of these components, what is the probability that exactly three would come from each plant?

Solution: There are many ways that 3 components could come from each plant, e.g., the sources of the 9 components could be ABBACCBAC. The possible ways correspond to rearrangements of the "word" AAABBBCCC, of which there are $\frac{9!}{3!3!3!} = 1680$.

Therefore the probability that exactly 3 components would come from each plant is

$$1680P(AAABBBCCC) = 1680(.5)^3(.3)^3(.2)^3 = .04536$$

Multinomial Distribution

The previous problem can be generalized:

Suppose a sequence of independent random variables $Y_1, \dots, Y_n$ each have r possible values with probabilities $p_1, \dots, p_r$. Let X_i be the number of $Y_1, \dots, Y_n$ which have value i. Then $(X_1, \dots, X_r)$ have a **multinomial** distribution. The joint pmf is

$$f(x_1, \dots, x_r) = \begin{cases} \frac{n!}{x_1! \cdots x_r!} p_1^{x_1} \cdots p_r^{x_r}, & \text{if } x_1 + \cdots + x_r = n \\ 0, & \text{otherwise} \end{cases}$$

When there are only two possible outcomes ($r = 2$), this reduces to a binomial distribution:

$$f(x, n-x) = \frac{n!}{x!(n-x)!} p_1^x p_2^{n-x} = \binom{n}{x} p_1^x (1-p_1)^{n-x}$$

Joint Probability Density

Random variables X and Y are said to have a **joint probability density** (joint pdf) $f(x, y)$ if

$$P(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f(x, y)\ dy\ dx$$

for all real constants $a \le b$ and $c \le d$.

To be a valid joint pdf, we must have $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\ dy\ dx = 1$.

Example

A bank operates a drive-up facility and a walk-up window. On a random day, let X be the proportion of time the drive-up facility is in use, and let Y be the proportion of time the walk-up window is in use. Check that a valid joint pdf for X and Y is given by

$$f(x,y) = \begin{cases} \frac{6}{5}(x+y^2), & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\ dy\ dx &= \int_0^1 \int_0^1 \frac{6}{5}(x+y^2)\ dy\ dx \\ &= \frac{6}{5}\int_0^1 \left(xy + \frac{y^3}{3}\right)\Big|_{y=0}^{1}\ dx \\ &= \frac{6}{5}\int_0^1 \left(x + \frac{1}{3}\right)\ dx = \frac{6}{5}\left(\frac{1}{2}+\frac{1}{3}\right) = 1 \end{aligned}$$

Marginal Probability Density Function

Let X and Y be random variables with joint pdf $f(x, y)$. The **marginal probability density function** (marginal pdf) of X is defined as

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\ dy$$

Similarly, the marginal pdf of Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\ dx$$

Note: f_X is simply the pdf of X considered as a random variable on its own.

X and Y are independent if their joint pdf is the product of the marginal pdfs:

$$f(x, y) = f_X(x) f_Y(y)$$

Example

In the bank example, the joint pdf of two random variables X and Y was given by

$$f(x,y) = \begin{cases} \frac{6}{5}(x+y^2), & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

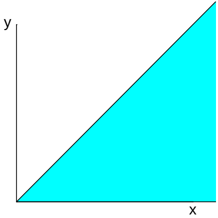
Find the marginal pdf of X.

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\ dy \\ &= \int_0^1 \frac{6}{5}(x+y^2)\ dy \\ &= \frac{6}{5}\left(xy + \frac{y^3}{3}\right)\Big|_{y=0}^{1} = \frac{6}{5}\left(x + \frac{1}{3}\right) \end{aligned}$$

Example

A system has two components; assuming their lifetimes X and Y are independent exponential random variables with $E(X) = 10$ and $E(Y) = 20$, what is the probability that the first component outlasts the second?

$$
\begin{aligned}
P(X \geq Y) &= \int_0^\infty \int_y^\infty f(x,y)\ dx\ dy \\
&= \int_0^\infty \int_y^\infty \frac{1}{200} e^{-\frac{x}{10} - \frac{y}{20}}\ dx\ dy \\
&= \frac{1}{200} \int_0^\infty e^{-y/20} \int_y^\infty e^{-x/10}\ dx\ dy \\
&= \frac{1}{200} \int_0^\infty e^{-y/20} \cdot 10e^{-y/10}\ dy \\
&= \frac{1}{20} \int_0^\infty e^{-\frac{3}{20}y}\ dy = \frac{1}{20} \cdot \frac{20}{3} = \frac{1}{3}
\end{aligned}
$$



Sum of Two Continuous Random Variables

Recall that for discrete random variables X and Y with joint pmf $f(x, y)$, their sum $X + Y$ has pmf

$$f_{X+Y}(t) = \sum_x f(x, t - x)$$

Given continuous random variables X and Y with joint pdf $f(x, y)$, their sum $X + Y$ has pdf

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f(x, t - x)\ dx$$

In particular, if X and Y are independent, then

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t - x)\ dx$$

Sum of Independent Normal Random Variables

With a bit of calculus and algebra, one may show:

If X and Y are independent normal random variables, then their sum $X + Y$ is also a normal random variable.

If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$, then

$$E(X+Y) = E(X) + E(Y) = \mu_1 + \mu_2$$
$$V(X+Y) = V(X) + V(Y) = \sigma_1^2 + \sigma_2^2$$

So $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Note: Since Y is normal, $-Y$ is normal, $-Y \sim N(-\mu_2, \sigma_2^2)$, so the difference $X - Y$ is also normal: $X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

Example

Suppose that the height of a randomly selected male adult in the US is normal with mean $\mu = 69.5$ inches and standard deviation $\sigma = 1.9$ inches. If two individuals from this population are selected at random, what is the probability that their height differs by more than 4 inches?

Solution: If the two individuals have heights X and Y, then $X, Y \sim N(69.5, 1.9^2)$, and so

$$X - Y \sim N(69.5 - 69.5, 1.9^2 + 1.9^2) = N(0, 7.22)$$

Thus $X - Y$ is normal with mean 0 and standard deviation $\sigma = \sqrt{7.22} \approx 2.687$.

$$\begin{aligned} P(|X - Y| > 4) &\approx P(|Z| > 4/2.687) \\ &\approx P(|Z| > 1.49) \\ &= 2P(Z < -1.49) = 2(.0681) = .1362 \end{aligned}$$

Covariance

Given random variables X and Y with means μ_X and μ_Y respectively, their **covariance** is

$$\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

As for variance, there is a shortcut formula for covariance:

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

The covariance has several useful properties:

1. $\mathrm{Cov}(X, X) = V(X)$
2. $\mathrm{Cov}(Y, X) = \mathrm{Cov}(X, Y)$
3. $\mathrm{Cov}(cX, Y) = c\,\mathrm{Cov}(X, Y)$
4. $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

The **correlation** between two random variables X and Y is

$$\rho = \mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Correlation is a measure of how strongly related two random variables are. It has several properties:

1. $-1 \leq \rho \leq 1$
2. If X and Y are independent, then $\mathrm{Corr}(X, Y) = 0$.
3. Changing the scale of X and/or Y does not affect the correlation, i.e. for any $c \neq 0$,

$$\mathrm{Corr}(cX, Y) = \mathrm{Corr}(X, Y) = \mathrm{Corr}(X, cY)$$

Example

In the bank example, the joint pdf of two random variables X and Y was given by

$$f(x,y) = \begin{cases} \frac{6}{5}(x+y^2), & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Find the correlation of X and Y.

The marginal pdfs of X and Y are, for $0 \le x \le 1, 0 \le y \le 1$,

$$f_X(x) = \int_0^1 \frac{6}{5}(x+y^2)\ dy = \frac{6}{5}\left(xy + \frac{y^3}{3}\right)\Big|_{y=0}^{1} = \frac{6}{5}\left(x + \frac{1}{3}\right)$$

$$f_Y(y) = \int_0^1 \frac{6}{5}(x+y^2)\ dx = \frac{6}{5}\left(\frac{x^2}{2} + xy^2\right)\Big|_{x=0}^{1} = \frac{6}{5}\left(\frac{1}{2} + y^2\right)$$

Example (continued)

From the marginal pdfs $f_X(x) = \frac{6}{5}(x + \frac{1}{3})$ and $f_Y(y) = \frac{6}{5}(\frac{1}{2} + y^2)$,

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\ dx = \frac{6}{5}\int_0^1 \left(x^2 + \frac{x}{3}\right) dx = \frac{6}{5}\left(\frac{1}{3} + \frac{1}{6}\right) = \frac{3}{5}$$

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y)\ dy = \frac{6}{5}\int_0^1 \left(\frac{y}{2} + y^3\right) dy = \frac{6}{5}\left(\frac{1}{4} + \frac{1}{4}\right) = \frac{3}{5}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\ dx = \frac{6}{5}\int_0^1 \left(x^3 + \frac{x^2}{3}\right) dx = \frac{6}{5}\left(\frac{1}{4} + \frac{1}{9}\right) = \frac{13}{30}$$

$$E(Y^2) = \int_{-\infty}^{\infty} y f_Y(y)\ dy = \frac{6}{5}\int_0^1 \left(\frac{y^2}{2} + y^4\right) dy = \frac{6}{5}\left(\frac{1}{6} + \frac{1}{5}\right) = \frac{11}{25}$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{13}{30} - \left(\frac{3}{5}\right)^2 = \frac{11}{150}$$

$$V(Y) = E(Y^2) - [E(Y)]^2 = \frac{11}{25} - \left(\frac{3}{5}\right)^2 = \frac{2}{25}$$

Example (continued)

Finally, from $f(x,y) = \frac{6}{5}(x+y^2)$ we calculate

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\ dx\ dy = \frac{6}{5}\int_0^1\int_0^1 (x^2y + xy^3)\ dx\ dy$$

$$= \frac{6}{5}\int_0^1 \left(\frac{x^3y}{3} + \frac{x^2y^3}{2}\right)\Big|_{x=0}^1\ dy$$

$$= \frac{6}{5}\int_0^1 \left(\frac{y}{3} + \frac{y^3}{2}\right)\ dy = \frac{6}{5}\left(\frac{1}{6} + \frac{1}{8}\right) = \frac{7}{20}$$

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{7}{20} - \frac{3}{5}\cdot\frac{3}{5} = \frac{-1}{100}$$

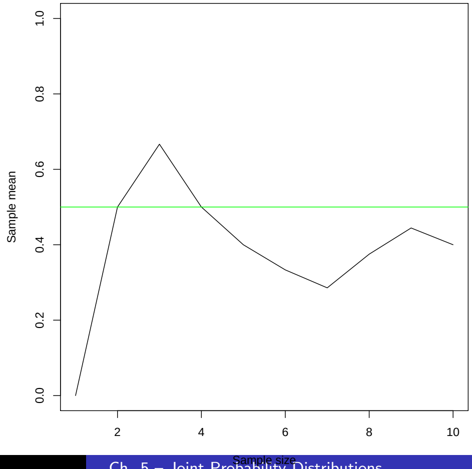
Therefore, from $V(X) = 11/150$ and $V(Y) = 2/25$,

$$\rho = \text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{-1/100}{\sqrt{11/150}\sqrt{2/25}} \approx -.131$$

Convergence of Sample Means

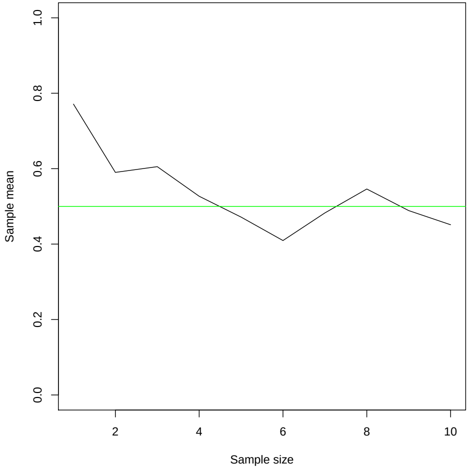
Suppose we generate a sequence of independent Bernoulli random variables $X_1, X_2, \ldots$ with $p = .5$ and take the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ using larger and larger sample sizes.

X_i	$\overline{X}$
0	0
1	1/2
1	2/3
0	2/4
0	2/5
0	2/6
0	2/7
1	3/8
1	4/9
0	4/10



Let's try doing the same thing using standard uniform random variables.

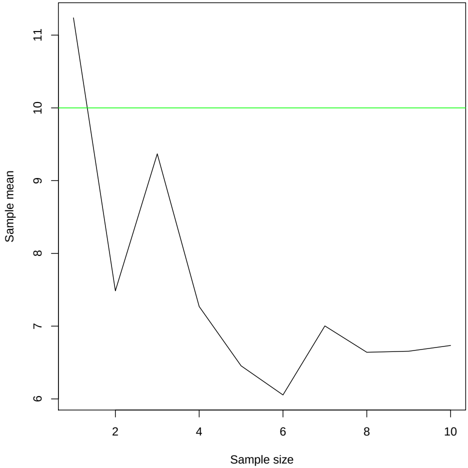
X_i	$\overline{X}$
0.771	0.771
0.409	0.590
0.636	0.605
0.291	0.527
0.251	0.472
0.099	0.409
0.922	0.483
0.990	0.546
0.031	0.489
0.113	0.451



Convergence of Sample Means

Let's try doing the same thing using exponential random variables with mean $\mu = 10$:

X_i	$\overline{X}$
11.24	11.24
3.73	7.49
13.13	9.37
0.97	7.27
3.19	6.46
4.05	6.05
12.69	7.00
4.11	6.64
6.77	6.66
7.45	6.73



A result in probability theory guarantees that this works no matter what the distribution of the X_i is, as long as $E(X_i)$ exists. For instance, the X_i could be Bernoulli, geometric, normal, hypergeometric, binomial, or just about anything else we could imagine.

Law of Large Numbers

Let $X_1, X_2, \ldots$ be independent, identically distributed (iid) random variables with mean μ. Then as $n \to \infty$, the sample mean $\overline{X}$ converges to μ with probability 1.

Sample Mean of iid Random Variables

Let $X_1, \ldots, X_n$ be independent, identically distributed (iid) random variables with mean μ and variance σ^2. Let $\overline{X}$ be the sample mean

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + \cdots + X_n)$$

We can compute the expected value and variance of $\overline{X}$:

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n}(n\mu) = \mu$$

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$$

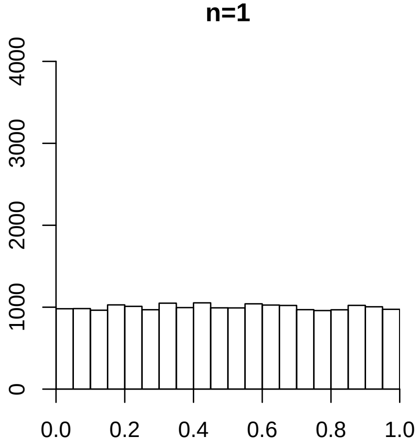
So the expected value of the sample mean $\overline{X}$ is μ, the same as the expected value of each summand. However, the standard deviation $\sigma/\sqrt{n}$ of $\overline{X}$ decreases as we increase the sample size n.

Simulate a standard uniform random variable 20000 times and draw the histogram.

Now simulate standard uniform random variables X_1 and X_2 and compute the sample mean $\overline{X} = \frac{1}{2}(X_1 + X_2)$; do this 20000 times and draw the histogram.

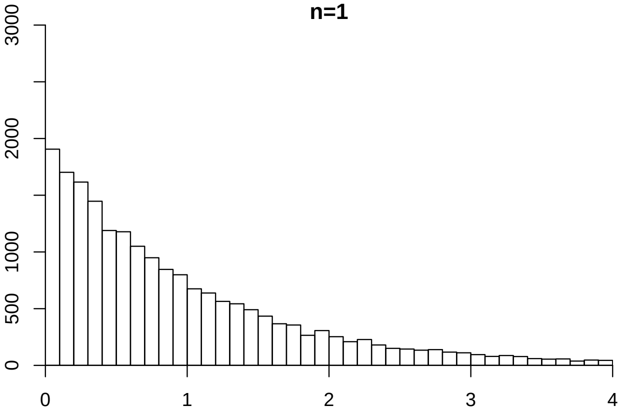
Now repeat this using larger sample sizes: draw the histogram of 20000 simulations of $\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$, for $n = 1, 2, 3, \ldots, 10$.

The histogram appears to approach a normal distribution.



Histogram of Sample Means

Now repeat this but instead of simulating standard uniform random variables, use exponential random variables with $\mu = 1$:



Central Limit Theorem

A powerful result of probability theory guarantees that in general, if n is large, then the sample mean $\overline{X}$ is approximately normal:

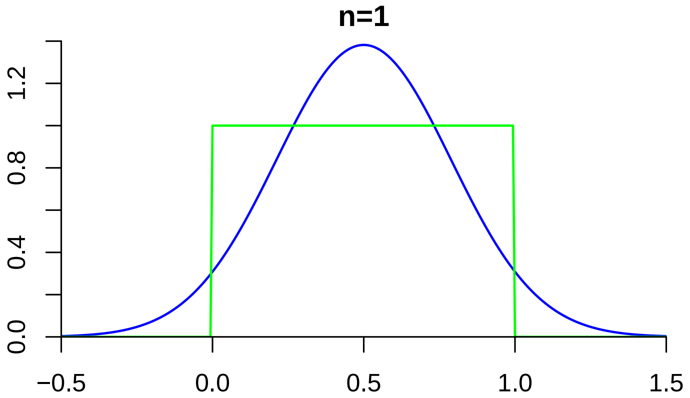
Let $X_1, X_2, \ldots$ be independent, identically distributed (iid) random variables with mean μ and standard deviation σ. Then

$$\lim_{n\to\infty} P\left(a \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq b\right) = P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$$

where Z is a standard normal random variable.

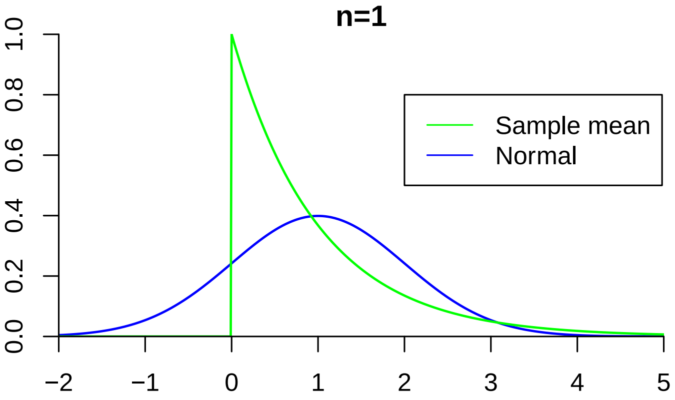
Distribution of Sample Means

Compare the pdf of a the sample mean $\overline{X}$ of standard uniform random variables with the pdf of a normal random variable as specified in the Central Limit Theorem:



Distribution of Sample Means

Compare the pdf of a the sample mean $\overline{X}$ of exponential random variables (with $\mu = 10$) with the pdf of a normal random variable as specified in the Central Limit Theorem:



Sums of Random Variables

Given iid random variables $X_1, \ldots, X_n$ with mean μ and standard deviation σ, the Central Limit Theorem tells us that if n is large, then the sample mean $\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$ is approximately normal.

It follows that the sum $X_1 + \cdots + X_n$ is also approximately normal, since the sum is the same as $\overline{X}$ scaled by a factor of n.

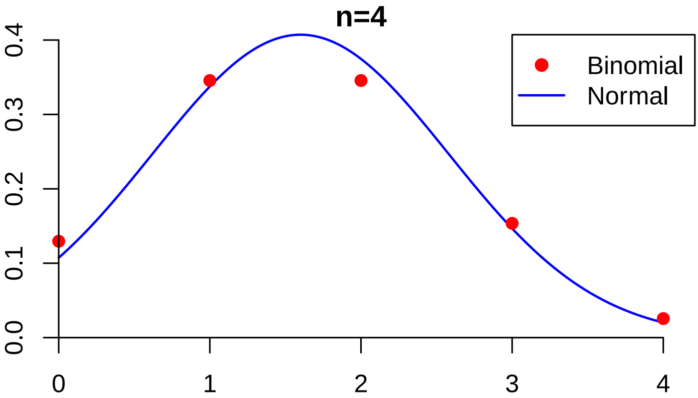
Examples:

- A binomial is a sum of iid Bernoulli random variables.
- A negative binomial is a sum of iid geometric random variables.
- A Poisson(μ) random variable, where μ is an integer, is a sum of iid Poisson(1) random variables.

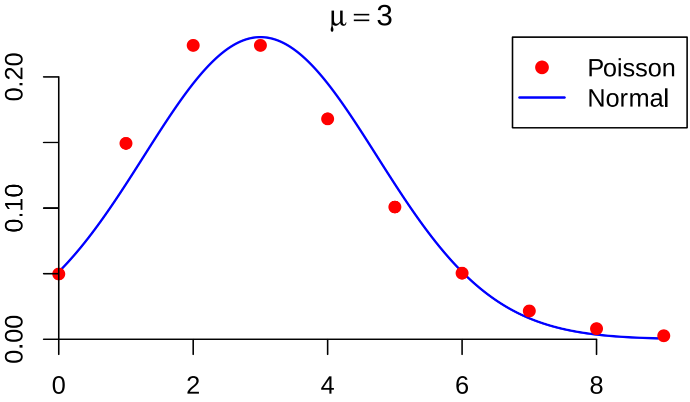
So all of these examples may be approximated as normal if n is large enough.

Normal Approximation to Binomial

Compare the pmf of a binomial random variable (with $p = .4$) to the pdf of the corresponding normal random variable:



Compare the pmf of a Poisson random variable to the pdf of the corresponding normal random variable:



Example

Bob is a candidate for political office in a large city and must take more than half the votes in order to win the election. Suppose a poll finds that 170 of 300 randomly sampled voters favor Bob. Approximate the probability that the poll would find so many voters favoring Bob if the true proportion were only .5.

Solution: The number X of sampled voters favoring Bob is a hypergeometric random variable; but since the population is large, we may approximate X as binomial, Bin(300, .5).

The Central Limit Theorem implies that X is approximately normal, with mean $\mu = np = 150$ and standard deviation $\sigma = \sqrt{np(1-p)} = \sqrt{75} \approx 8.66$.

$$\begin{aligned} P(X \geq 170) &\approx P(8.66Z + 150 \geq 170) \\ &\approx P(Z \geq 2.31) \\ &= P(Z \leq -2.31) = \Phi(-2.31) \approx .0104 \end{aligned}$$

Example

A Geiger counter placed next to a certain radioactive specimen clicks an average of 90 times per minute. Over a 10 minute period, approximate the probability that it would click 800 times or less.

Solution: The number of clicks X in a 10 minute period has a Poisson distribution with mean $\mu = 900$. The variance of X is $V(X) = \mu = 900$, so the standard deviation is $\sigma = \sqrt{900} = 30$.

Since μ is large, the distribution of X is approximately normal. So we can approximate X as $30Z + 900$ where Z is a standard normal random variable. Then

$$\begin{aligned} P(X \leq 800) &\approx P(30Z + 900 \leq 800) \\ &\approx P(Z \leq -3.33) = \Phi(-3.33) \approx .0004 \end{aligned}$$

Summary

- **Joint pmf** of discrete random variables X and Y:

$$f(x, y) = P(X = x \cap Y = y)$$

- **Joint pdf** of continuous random variables X and Y:

$$P(a \le X \le b \cap c \le Y \le d) = \int_a^b \int_c^d f(x, y)\ dy\ dx$$

- **Covariance** of two random variables X and Y:
  - Definition: $\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$
  - Shortcut formula: $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$
  - Properties:
    1. $\text{Cov}(X, X) = V(X)$
    2. $\text{Cov}(Y, X) = \text{Cov}(X, Y)$
    3. $\text{Cov}(cX, Y) = c\,\text{Cov}(X, Y)$
    4. $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$

Summary

- **Correlation**:
$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$
- **Sample Mean**:
$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$
  - If $E(X_i) = \mu$ and $V(X_i) = \sigma^2$, then
$$E(\overline{X}) = \mu$$
$$V(\overline{X}) = \frac{\sigma^2}{n}$$
- **Law of Large Numbers**: As the sample size n increases, $\overline{X}$ converges to $E(X)$.
- **Central Limit Theorem**: As the sample size n increases, $\overline{X}$ becomes approximately normal. More precisely, $\frac{\overline{X} - E(X)}{\sigma/\sqrt{n}}$ converges to a standard normal distribution.