MATH 321: Test 1 Study Guide

Lecture 14 – Bivariate Distributions (4.1 and 4.4)

Joint pmf and pdf

- Discrete definition: The joint pmf is defined as $f(x, y) = P(X = x, Y = y)$ for all $(x, y) \in \mathbb{R}^2$ and has properties
  1. $0 \leq f_{X,Y}(x, y) \leq 1 \quad$ for all x, y
  2. $\displaystyle\sum_x \sum_y f(x, y) = \sum_y \sum_x f(x, y) = 1$
  3. Let A be any subset of $\mathbb{R}^2$, then $P((X, Y) \in A) = \sum \sum_A f(x, y)$
- Continuous definition: The joint pdf is a function $f(x, y)$ from $\mathbb{R}^2$ into $\mathbb{R}$ such that
  1. $f_{X,Y}(x, y) \geq 0 \quad$ for all x, y
  2. $\displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y\,\mathrm{d}x = 1$
  3. For $A \subset \mathbb{R}^2$, $\displaystyle P((X, Y) \in A) = \int \int_A f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int \int_A f(x, y)\,\mathrm{d}y\,\mathrm{d}x$

Marginal distributions

- Discrete definition: Let (X, Y) have joint pmf $f(x, y)$. Then, the marginal pmfs are given by

  $$f_X(x) = \sum_y f_{X,Y}(x, y) \qquad \text{and} \qquad f_Y(y) = \sum_x f_{X,Y}(x, y)$$

- Continuous definition: Let (X, Y) have joint pdf $f(x, y)$. Then the marginal pdfs are defined by:

  $$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y \qquad \text{and} \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x$$

Expected values of a function of a random variable

- Definition: Let $g(X, Y)$ be a function of a bivariate random vector (X, Y).
  (a) If X and Y are discrete with joint pmf $f(x, y)$,

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

  (b) If X and Y are continuous with joint pdf $f(x, y)$,

  $$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

Special expectations

- Definitions: Let (X_1, X_2) be a bivariate random vector with joint pmf / pdf $f(x_1, x_2)$.
  i) If $g(X_1, X_2) = X_1$, then $E[g(X_1, X_2)] = E(X_1) = \mu_{X_1}$
  ii) If $g(X_1, X_2) = (X_1 - \mu_1)^2$, then $E[g(X_1, X_2)] = E[(X_1 - \mu_1)^2] = \sigma^2_{X_1}$
  iii) If $g(X_1, X_2) = \mathrm{e}^{tX_1}$, then $E[g(X_1, X_2)] = E(\mathrm{e}^{tX_1}) = M_{X_1}(t)$

Expected value of $X + Y$ and XY

- Theorem: Expected value of a sum of two random variables

  If $g(X, Y) = X + Y$, then $E(X + Y) = E(X) + E(Y)$
- Generalized theorem: If $g_1(X, Y)$ and $g_2(X, Y)$ are two functions and a, b and c are constants, then

  $E[ag_1(X, Y) + bg_2(X, Y) + c] = aE[g_1(X, Y)] + bE[g_2(X, Y)] + c$
- Theorem: Expected value of a product of two random variables

  If $g(X, Y) = XY$ and $X \perp\!\!\!\perp Y$, then $E(XY) = E(X) \cdot E(Y)$

Lecture 15 – Conditional Distributions (4.3)

Conditional pmf / pdf

- Definition: Let (X, Y) be a bivariate random vector with joint pmf / pdf $f(x, y)$ and marginal pmfs / pdfs $f_X(x)$ and $f_Y(y)$.
  (a) Given x such that $f_X(x) > 0$, $\qquad f(y \mid x) = \dfrac{f(x, y)}{f_X(x)}$
  (b) Given y such that $f_Y(y) > 0$, $\qquad f(x \mid y) = \dfrac{f(x, y)}{f_Y(y)}$

Probabilities

- For $A \subset \mathbb{R}^2$,

  Discrete: $P(X \in A \mid Y = y) = \displaystyle\sum_{x \in A} P(X = x \mid Y = y) = \sum_{x \in A} f(x \mid y)$

  Continuous: $P(X \in A \mid Y = y) = \displaystyle\int_A f(x \mid y)\,\mathrm{d}x$

Relationship between joint pmf and conditional pmfs

- Theorem: For bivariate random vector (X, Y) with joint pmf / pdf $f(x, y)$ and x and y such that $f_X(x) > 0$ and $f_Y(y) > 0$,

  $f(x, y) = f_Y(y) \cdot f(x \mid y) = f_X(x) \cdot f(y \mid x)$

Conditional expected values

- Definition: Let $g(Y)$ be a function of Y, then the conditional expected value of $g(Y)$ given that $X = x$ is given by

  $$E[g(Y) \mid x] = \sum_y g(y) f(y \mid x) \qquad \text{and} \qquad E[g(Y) \mid x] = \int_{-\infty}^{\infty} g(y) f(y \mid x)\, \mathrm{d}y$$

- Conditional mean and variance definitions (assuming X and Y are discrete):

  i) If $g(Y) = Y$, then the conditional mean of Y given $X = x$ is

  $$E(Y \mid X = x) = \sum_y y\, f(y \mid x) = \mu_{Y|X}$$

  ii) If $g(Y) = (Y - \mu_{Y|X})^2$, then the conditional variance of Y given $X = x$ is

  $$E[(Y - \mu_{Y|X})^2 \mid X = x] = \sum_y (y - \mu_{Y|X})^2\, f(y \mid x) = \sigma^2_{Y|X}$$

<u>**Lecture 16 – Independence and the Correlation Coefficient**</u> (4.1, 4.2, and 4.4)

Independence for random variables

- Definition: Let (X, Y) be a bivariate random vector with joint pdf / pmf $f(x, y)$ and marginal pdfs / pmfs $f_X(x)$ and $f_Y(y)$. Then X and Y are called independent random variables if, for every $x \in \mathbb{R}$ and $y \in \mathbb{R}$,

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

- Checking independence theorem: X and Y are independent random variables if and only if

  $$f(x, y) = g(x) \cdot h(y), \qquad a \le x \le b,\ c \le y \le d,$$

  where $g(x)$ is a nonnegative function of x alone and $h(y)$ is a nonnegative function of y alone

Conditional distributions and independence

- Theorem: If X and Y are independent, $f(x \mid y) = f_X(x) \qquad \text{and} \qquad f(y \mid x) = f_Y(y)$

Using independence

- Theorem: Let X and Y be independent random variables.

  (a) For any $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$, $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$

  (b) Let $g(x)$ be a function only of x and $h(y)$ be a function only of y. Then

  $$E[g(X) \cdot h(Y)] = E[g(X)] \cdot E[h(Y)]$$

Definition, theorems and properties of covariance

- Definition: The covariance of X and Y is the number defined by: $\mathrm{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$
- If (X,Y) is discrete, then $E[(X-\mu_X)(Y-\mu_Y)] = \sum_x \sum_y (x-\mu_x)(y-\mu_y\, f(x,y)$
- Alternate calculation for covariance: $\mathrm{Cov}(X,Y) = E(XY) - E(X)\cdot E(Y)$
- Variance is a special case of covariance: $V(X) = \mathrm{Cov}(X,X)$
- Order in covariance does not matter (i.e. symmetric): $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
- Covariance of a random variable and a constant is zero: If c is a constant, then $\mathrm{Cov}(X,c) = 0$
- Can factor out coefficients in covariance: $\mathrm{Cov}(aX,bY) = ab\cdot\mathrm{Cov}(X,Y)$
- Can factor out coefficients, but added constants disappear: $\mathrm{Cov}(aX+c, bY+d) = ab\cdot\mathrm{Cov}(X,Y)$
- Distributive property of covariance: $\mathrm{Cov}(X,Y+Z) = \mathrm{Cov}(X,Y) + \mathrm{Cov}(X,Z)$
- Independence and covariance theorem: If $X \perp\!\!\!\perp Y$ then $\mathrm{Cov}(X,Y) = 0$

Correlation definition and properties

- Definition: $\rho_{XY} = \mathrm{Corr}(X,Y) = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)V(Y)}} = \dfrac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$
- Theorem: For any random variable X and Y,
  i) $-1 \le \rho_{XY} \le 1$
  ii) $\rho_{XY} = 1$ if and only if there exist numbers $a > 0$ and b such that $P(Y = aX + b) = 1$.
  iii) $\rho_{XY} = -1$ if and only if there exist numbers $a < 0$ and b such that $P(Y = aX + b) = 1$.
  iv) When $\rho_{XY} = 0$, X and Y are uncorrelated.

Variance of $X+Y$

- Theorem: Variance of a sum of two random variables

  $V(X+Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X,Y)$

  If $X \perp\!\!\!\perp Y$, then $V(X+Y) = V(X) + V(Y)$

Lecture 17 – Several Random Variables (5.3 and 5.4)

Definitions and theorems

- Joint distributions
  - Discrete definition: If $\mathbf{X} = (X_1, \ldots, X_n)$ a discrete random vector (the range is countable), then the joint pmf of $\mathbf{X}$ is the function defined by

    $f(\mathbf{x}) = f(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n)$ for each $(x_1, \ldots, x_n) \in \mathbb{R}^n$

    Then for any $A \subset \mathbb{R}^n$,

    $$P(\mathbf{X} \in A) = \sum_{\mathbf{x} \in A} f(\mathbf{x})$$
  - Continuous definition: If $\mathbf{X} = (X_1, \ldots, X_n)$ a continuous random vector, then the joint pdf of $\mathbf{X}$ is the function $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ that satisfies

    $$P(\mathbf{X} \in A) = \int \cdots \int_A f(\mathbf{x})\, \mathrm{d}\mathbf{x} = \int \cdots \int_A f(x_1, \ldots, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_n$$
- Expected values: Let $g(\mathbf{x})$ be a real-valued function defined on the range of $\mathbf{X}$. The expected value of $g(\mathbf{X})$ is

  $$E[g(\mathbf{X})] = \quad \overbrace{\sum_{\mathbf{x} \in \mathbb{R}^n} g(\mathbf{x}) f(\mathbf{x})}^{\text{Discrete}} \qquad \overbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\mathbf{x}) f(\mathbf{x}) \mathrm{d}x_1 \cdots \mathrm{d}x_n}^{\text{Continuous}}$$
- Marginal distributions: The marginal pdf or pmf of any subset of the coordinates of $(X_1, \ldots, X_n)$ can be computed by integrating or summing the joint pdf or pmf over all possible values of the other coordinates.
- Conditional distributions: The conditional pmf or pdf of a subset of the coordinates of $(X_1, \ldots, X_n)$ given the value of the remaining coordinates is obtained by dividing the joint pdf or pmf by the marginal pdf or pmf of the remaining coordinates.

Independence

- Definition: Let random variables $X_1, \ldots, X_n$ have joint pdf (or pmf) $f(x_1, \ldots, x_n)$ and let $f_{X_i}(x_i)$ be the marginal pdf (or pmf) of X_i. Then $X_1, \ldots, X_n$ are mutually independent random variables if, for every $(x_1, \ldots, x_n)$, the joint pdf (or pmf) can be written as

  $$f(X_1, \ldots, X_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$
- Conditional distributions: If $X_1, \ldots, X_n$ are mutually independent, the conditional distribution of any subset of the coordinates, given the values of the rest of the coordinates, is the same as the marginal distribution of the subset.
- Expected value: Let $X_1, \ldots, X_n$ be mutually independent random variables. Let $g_1, \ldots, g_n$ be real-valued functions such that $g_i(x)$ is a function only of x_i, $i = 1, \ldots, n$. Then

  $$E[g_1(X_1) \cdots g_n(X_n)] = \prod_{i=1}^{n} E[g_i(x_i)]$$

Linear functions of random variables

- Definition: A linear function of random variables consists of n random variables $X_1, \ldots, X_n$ and n coefficient $a_1, \ldots, a_n$

  $$a_1X_1 + a_2X_2 + \cdots + a_nX_n = \sum_{i=1}^{n} a_iX_i$$

- Expected value of a linear function of random variables

  $$E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n)$$

- Variance of a linear function of random variables

  $$V[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = \sum_{i=1}^{n} a_i^2 V(X_i) + 2\sum_{i<j} a_ia_j \mathrm{Cov}(X_i, X_j)$$

  If $X_1, \ldots, X_n$ are mutually independent,

  $$V[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = \sum_{i=1}^{n} a_i^2 V(X_i)$$

Mgf of sums of independent random variables

- Theorem: Let $X_1, \ldots, X_n$ be mutually independent random variables with mgfs $M_{X_1}(t), \ldots, M_{X_n}(t)$. Let $Y = X_1 + \cdots + X_n$.

  $$M_Y(t) = M_{X_1+\cdots+X_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

  If $X_1, \ldots, X_n$ all have the same distribution with mgf $M_X(t)$, then

  $$M_Y(t) = \big[M_X(t)\big]^n$$

Sums of linear combinations of random variables

- Theorem: Let $X_1, \ldots, X_n$ be mutually independent random variables with mgfs $M_{X_1}(t), \ldots, M_{X_n}(t)$. Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be fixed constants. Let $Y = (a_1X_1 + b_1) + \cdots + (a_nX_n + b_n)$. Then the mgf of Y is

  $$M_Y(t) = \big(\mathrm{e}^{t\sum b_i}\big) M_{X_1}(a_1t) \cdots M_{X_n}(a_nt)$$

- Sum of linear function of normals theorem: Let $X_1, \ldots, X_n$ be mutually independent random variables with $X_i \sim \text{Normal}\,(\mu_i, \sigma_i^2)$. Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be fixed constants. Then,

  $$Y = \sum_{i=1}^{n}(a_iX_i + b_i) \sim \text{Normal}\left(\mu = \sum_{i=1}^{n}(a_i\mu_i + b_i),\, \sigma^2 = \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$$

Distributions

Discrete Distributions	
Discrete uniform (N_0, N_1)	
Pmf	$P(X = x \mid N_0, N_1) = \frac{1}{N_1 - N_0 + 1}; \qquad x = N_0, \ldots, N_1; \qquad N_0 \leq N_1$
Mean and Variance	$E(X) = \frac{N_0 + N_1}{2}, \qquad V(X) = \frac{(N_1 - N_0 + 1)^2 - 1}{12}$
Mgf	$M_X(t) = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} \mathrm{e}^{tx}$
Notes	
Bernoulli (p)	
Pmf	$P(X = x \mid p) = p^x (1-p)^{1-x}; \qquad x = 0, 1; \qquad 0 < \mathrm{p} < 1$
Mean and Variance	$E(X) = p, \qquad V(X) = p(1-p) = pq$
Mgf	$M_X(t) = (1-p) + p\mathrm{e}^t = q + p\mathrm{e}^t$
Notes	Special case of binomial with $n = 1$.
Binomial (n, p)	
Pmf	$P(X = x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}; \qquad x = 0, 1, \ldots, n; \qquad 0 < p < 1$
Mean and Variance	$E(X) = np, \qquad V(X) = np(1-p) = npq$
Mgf	$M_X(t) = (q + p\mathrm{e}^t)^n$
Notes	Sum of *iid* bernoulli RVs.
Geometric (p)	
Pmf	$P(X = x \mid p) = q^{x-1} p; \qquad x = 1, 2, \ldots; \qquad 0 < p < 1$
Cdf	$F_X(x \mid p) = 1 - q^x$
Mean and Variance	$E(X) = \frac{1}{p}, \qquad V(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$
Mgf	$M_X(t) = \frac{p\mathrm{e}^t}{1 - q\mathrm{e}^t}; \qquad t < -\ln(q)$
Notes	Special case of negative binomial with $r = 1$. * See other geometric probabilities. Alternate form $Y = X - 1$. This distribution is *memoryless*: $P(X > s \mid X > t) = P(X > s - t); \qquad s > t$.

Negative binomial (r, p)	
Pmf	$P(X = x \mid r, p) = P(X = x \mid r, p) = \binom{x-1}{r-1} p^r q^{x-r}; \qquad x = r, r+1, \ldots; \qquad 0 < p < 1$
Mean and Variance	$E(X) = \frac{r}{p}, \qquad V(X) = \frac{r(1-p)}{p^2} = \frac{rq}{p^2}$
Mgf	$M_X(t) = \left[\frac{pe^t}{1-qe^t}\right]^r; \qquad t < -\ln(q)$
Notes	Sum of *iid* geometric RVs.
Hypergeometric (N, M, K)	
Pmf	$P(X = x \mid r, p) = P(X = x \mid N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}; \qquad x = 0, 1, \ldots, \min(M, K)$
Mean and Variance	$E(X) = K\left(\frac{M}{N}\right), \qquad V(X) = K\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-K}{N-1}\right)$
Mgf	
Notes	If do not require $M \geq K$, $\mathcal{X} = \{\max(0, K + M - N), \ldots, \min(M, K)\}$, mean and variance converge to that of binomial ($n = K, p = M/K$) when $N \to \infty$.
Poisson (λ)	
Pmf	$P(X = x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots; \qquad \lambda > 0$
Mean and Variance	$E(X) = \lambda, \qquad V(X) = \lambda$
Mgf	$M_X(t) = e^{\lambda(e^t - 1)}$
Notes	If $X_i \overset{\perp\!\!\!\perp}{\sim}$ Poisson (λ_i), then $\sum X_i \sim \text{Poisson}(\lambda = \sum \lambda_i)$.

Other geometric probabilities

- Let $X \sim \text{Geometric}(p)$.

$$\begin{aligned}
P(X < \infty) &= 1 \\
P(X > x) &= q^x \\
P(X \geq x) &= q^{x-1} \\
P(a < X \leq b) &= q^a - q^b \\
P(a \leq X \leq b) &= q^{a-1} - q^b
\end{aligned}$$

Continuous Distributions

Continuous uniform (a, b)

Pdf	$f(x \mid a,b) = \frac{1}{b-a}, \qquad a \le x \le b; \qquad a,b \in \mathbb{R}, \qquad a \le b$
Cdf	$F(x) = \frac{x-a}{b-a} \qquad a \le x \le b$
Survival	$S(t) = \frac{b-t}{b-a} \qquad a \le t \le b \qquad$ if $T \sim \text{Uniform}\,(a,b)$
Mean and Variance	$E(X) = \frac{a+b}{2}; \qquad V(X) = \frac{(b-a)^2}{12}$
Mgf	$M_X(t) = \frac{\mathrm{e}^{tb} - \mathrm{e}^{ta}}{t(b-a)} \qquad t \neq 0$
Notes	

Exponential (λ)

Pdf	$f(t \mid \lambda) = \lambda \mathrm{e}^{-\lambda t}, \qquad t \ge 0; \qquad \lambda > 0$
Cdf	$F(t) = 1 - \mathrm{e}^{-\lambda t} \qquad t \ge 0$
Survival	$S(t) = \mathrm{e}^{-\lambda t} \qquad t \ge 0$
Mean and Variance	$E(X) = \frac{1}{\lambda}; \qquad V(X) = \frac{1}{\lambda^2}$
Mgf	$M_X(t) = \frac{\beta}{\beta - t} \qquad t < \beta; \qquad$ if $T \sim \text{Exp}\,(\beta)$
Notes	Special case of gamma with $\alpha = 1, \beta$. This distribution is *memoryless*: $P(T > a+b \mid T > a) = P(T > b); \qquad a,b > 0$. Rate parameterization is given; alternate parameterization is with scale $\theta = 1/\lambda$.

Gamma (α, β)

Pdf	$f(x \mid \alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, \mathrm{e}^{-\beta x}, \qquad x \ge 0; \qquad \alpha, \beta > 0$
Cdf	N/A
Mean and Variance	$E(X) = \frac{\alpha}{\beta} \qquad V(X) = \frac{\alpha}{\beta^2}$
Mgf	$M_X(t) = \left(\frac{\beta}{\beta - t}\right)^\alpha \qquad t < \beta$
Notes	$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}\, \mathrm{e}^{-x}\, \mathrm{d}x$ Sum of *iid* exponential RVs. A special case is exponential $(\alpha = 1, \beta)$. Rate parameterization is given; alternate parameterization is with scale $\theta = 1/\beta$.

Normal (μ, σ^2)

Pdf	$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\, \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty; \qquad -\infty < \mu < \infty, \quad \sigma > 0$
Cdf	N/A
Mean and Variance	$E(X) = \mu, \qquad V(X) = \sigma^2$
Mgf	$M_X(t) = \exp\left[\mu t + \frac{\sigma^2 t^2}{2}\right]$
Notes	Special case: Standard normal $Z \sim \text{Normal}\,(\mu = 0, \sigma^2 = 1)$.

Lognormal (μ, σ^2)

Pdf	$f(y \mid \mu, \sigma^2) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln(y)-\mu)^2}{2\sigma^2}\right]; \qquad y \geq 0; \qquad -\infty < \mu < \infty; \qquad \sigma > 0$
Mean and Variance	$E(Y) = e^{\mu + \frac{\sigma^2}{2}}, \qquad V(Y) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$
Mgf	
Notes	If $Y \sim$ Lognormal $\Longrightarrow \ln(Y) \sim$ Normal (μ, σ^2); equivalently, if $X \sim$ Normal (μ, σ^2) and $Y = e^X \Longrightarrow Y \sim$ Lognormal. μ and σ^2 represent the mean and variance of the normal random variable X which appears in the exponent.

Beta (α, β)

Pdf	$f(x \mid \alpha, \beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}; \qquad 0 \leq x \leq 1; \qquad \alpha, \beta > 0$
Mean and Variance	$E(X) = \frac{\alpha}{\alpha+\beta}, \qquad V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Mgf	
Notes	$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$