Recap

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$
  - Relation between P_X and p_X

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$
  - Relation between P_X and p_X
    $p_X(a) := P_X(\{a\}) = \mathbb{P}\{\omega \in \Omega : X(\omega) = a\}$

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$
  - Relation between P_X and p_X
    $p_X(a) := P_X(\{a\}) = \mathbb{P}\{\omega \in \Omega : X(\omega) = a\}$
- Continuous variables and relation between $\mathbb{P}, P_X, F_X, f_X$

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$
  - Relation between P_X and p_X
    $p_X(a) := P_X(\{a\}) = \mathbb{P}\{\omega \in \Omega : X(\omega) = a\}$
- Continuous variables and relation between $\mathbb{P}, P_X, F_X, f_X$
  - Relation between f_X and F_X is $F_X(a) = \int_{-\infty}^{a} f_X(x)dx$.

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$
  - Relation between P_X and p_X
    $p_X(a) := P_X(\{a\}) = \mathbb{P}\{\omega \in \Omega : X(\omega) = a\}$
- Continuous variables and relation between $\mathbb{P}, P_X, F_X, f_X$
  - Relation between f_X and F_X is $F_X(a) = \int_{-\infty}^{a} f_X(x)dx$.
  - $\frac{dF_X(x)}{dx} = f_X(x)$ or $P_X(x < X \leq x + h) \simeq f_X(x)h$.

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$
  - Relation between P_X and p_X
    $p_X(a) := P_X(\{a\}) = \mathbb{P}\{\omega \in \Omega : X(\omega) = a\}$
- Continuous variables and relation between $\mathbb{P}, P_X, F_X, f_X$
  - Relation between f_X and F_X is $F_X(a) = \int_{-\infty}^{a} f_X(x)dx$.
  - $\frac{dF_X(x)}{dx} = f_X(x)$ or $P_X(x < X \leq x + h) \simeq f_X(x)h$.
- Mean, Variance, Moments, $E[g(X)]$, Linearity & Examples

Recap

- Discrete random variables and relation between $\mathbb{P}, P_X, F_X, p_X$.
  - Relation between p_X and F_X
    $F_X(a) = \sum_{x \leq a} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq a\}$.
  - Relation between P_X and F_X
    $F_X(a) := P_X((-\infty, a]) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in (-\infty, a]\}$
  - Relation between P_X and p_X
    $p_X(a) := P_X(\{a\}) = \mathbb{P}\{\omega \in \Omega : X(\omega) = a\}$
- Continuous variables and relation between $\mathbb{P}, P_X, F_X, f_X$
  - Relation between f_X and F_X is $F_X(a) = \int_{-\infty}^{a} f_X(x)dx$.
  - $\frac{dF_X(x)}{dx} = f_X(x)$ or $P_X(x < X \leq x + h) \simeq f_X(x)h$.
- Mean, Variance, Moments, $E[g(X)]$, Linearity & Examples

$F_X : \mathbb{R} \to [0, 1]$ is non-decreasing and right continuous.

Gaussian random variable ($\mathcal{N}(\mu, \sigma^2)$)

Gaussian random variable ($\mathcal{N}(\mu, \sigma^2)$)

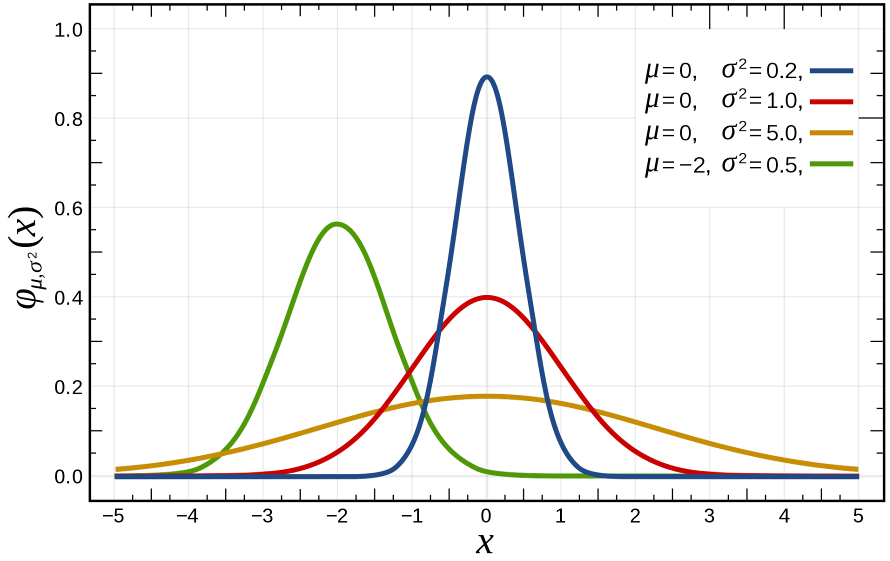
- This is a real valued r.v. with two paramerters, μ and σ.

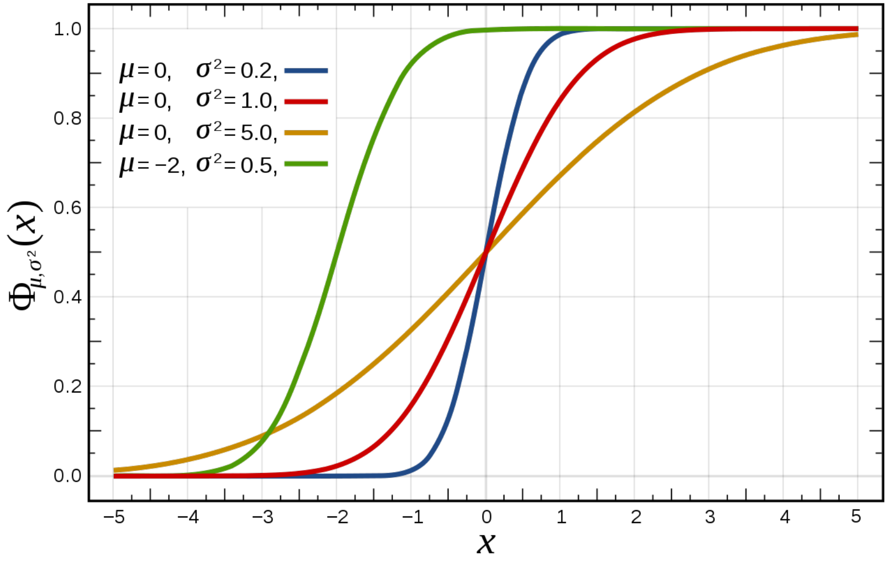
Gaussian random variable ($\mathcal{N}(\mu, \sigma^2)$)

- This is a real valued r.v. with two paramerters, μ and σ.
- Its pdf $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ for all $x \in \mathbb{R}$.

Gaussian random variable ($\mathcal{N}(\mu, \sigma^2)$)

- This is a real valued r.v. with two paramerters, μ and σ.
- Its pdf $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ for all $x \in \mathbb{R}$.
- Verify: $\int_{-\infty}^{\infty} f_X(x)dx$ $E[X] = \mu$ and $Var(X) = \sigma^2$.





Standard Normal random variable ($\mathcal{N}(0, 1)$)

Standard Normal random variable ($\mathcal{N}(0, 1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$?

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this?

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this? $(= \sqrt{2\pi})$

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this? $(= \sqrt{2\pi})$
- The CDF of standard normal, denoted by $\Phi(x)$ is given by

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this? $(= \sqrt{2\pi})$
- The CDF of standard normal, denoted by $\Phi(x)$ is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this? $(= \sqrt{2\pi})$
- The CDF of standard normal, denoted by $\Phi(x)$ is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

- $Q(x) := 1 - \Phi(x)$ is the Complimentary CDF $(P(X > x))$.

Standard Normal random variable ($\mathcal{N}(0, 1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this? $(= \sqrt{2\pi})$
- The CDF of standard normal, denoted by $\Phi(x)$ is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

- $Q(x) := 1 - \Phi(x)$ is the Complimentary CDF $(P(X > x))$.
  A closely related cousin in the error function
  $erf(x) = \frac{2}{\sqrt{pi}} \int_0^x e^{t^2} dt$.

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this? $(= \sqrt{2\pi})$
- The CDF of standard normal, denoted by $\Phi(x)$ is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

- $Q(x) := 1 - \Phi(x)$ is the Complimentary CDF ($P(X > x)$).
  A closely related cousin in the error function
  $erf(x) = \frac{2}{\sqrt{pi}} \int_0^x e^{t^2} dt$.
- $\Phi =$These values are recorded in a table. (Fig. 3.10 in Bertsekas)

Standard Normal random variable ($\mathcal{N}(0,1)$)

- When $\mu = 0$ and $\sigma = 1$, it is called as a standard normal.
- In this case $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.
- What is $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx$? How do you even solve this? $(= \sqrt{2\pi})$
- The CDF of standard normal, denoted by $\Phi(x)$ is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

- $Q(x) := 1 - \Phi(x)$ is the Complimentary CDF $(P(X > x))$. A closely related cousin in the error function $erf(x) = \frac{2}{\sqrt{pi}} \int_0^x e^{t^2} dt$.
- $\Phi =$These values are recorded in a table. (Fig. 3.10 in Bertsekas)
- https://en.wikipedia.org/wiki/Gaussian_function

Normality preserved under Linear Transformations

Normality preserved under Linear Transformations

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b$ is also a normal variable with $E[Y] = a\mu + b$ and variance $a^2\sigma^2$. (To be proved later)

Normality preserved under Linear Transformations

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b$ is also a normal variable with $E[Y] = a\mu + b$ and variance $a^2\sigma^2$. (To be proved later)

- Suppose X is standard normal, then find a and b such that $Y \sim \mathcal{N}(\mu, \sigma^2)$

Normality preserved under Linear Transformations

> If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b$ is also a normal variable with $E[Y] = a\mu + b$ and variance $a^2\sigma^2$. (To be proved later)

- Suppose X is standard normal, then find a and b such that $Y \sim \mathcal{N}(\mu, \sigma^2)$
- In this case, the CDF of Y in terms of X is given by $\Phi(\frac{x-\mu}{\sigma})$.

Significance of Gaussian r.v.

Significance of Gaussian r.v.

- Key role in Central limit theorem.

Significance of Gaussian r.v.

- Key role in Central limit theorem.
- $\frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \dfrac{\sigma^2}{n})$ where X_i is any random variable with mean μ and variance σ^2.

Significance of Gaussian r.v.

- Key role in Central limit theorem.
- $\frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ where X_i is any random variable with mean μ and variance σ^2.
- Building block for multinomial Gaussian vector and Gaussian processes (GP).

Significance of Gaussian r.v.

- Key role in Central limit theorem.
- $\frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ where X_i is any random variable with mean μ and variance σ^2.
- Building block for multinomial Gaussian vector and Gaussian processes (GP).
- Gaussian process are used in Bayesian Optimization (black-box optimization).

Significance of Gaussian r.v.

- Key role in Central limit theorem.
- $\frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ where X_i is any random variable with mean μ and variance σ^2.
- Building block for multinomial Gaussian vector and Gaussian processes (GP).
- Gaussian process are used in Bayesian Optimization (black-box optimization).
- Brownian motion is a type of GP and is used in Finance.

Significance of Gaussian r.v.

- Key role in Central limit theorem.
- $\frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \dfrac{\sigma^2}{n})$ where X_i is any random variable with mean μ and variance σ^2.
- Building block for multinomial Gaussian vector and Gaussian processes (GP).
- Gaussian process are used in Bayesian Optimization (black-box optimization).
- Brownian motion is a type of GP and is used in Finance.
- GP Regression, Gaussian mixture models, used widely in ML.

List of Probability distributions ...

List of Probability distributions ...

https://en.wikipedia.org/wiki/List_of_probability_distributions

List of Probability distributions ...

https://en.wikipedia.org/wiki/List_of_probability_distributions

Important ones are Beta, Gamma, Erlang, Logistic, Weibull ....

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.
- If X is discrete, $M_X(t) = \sum_{x \in \Omega'} e^{tx} p_X(x)$.

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.
- If X is discrete, $M_X(t) = \sum_{x \in \Omega'} e^{tx} p_X(x)$.
- If X is continuous, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$.

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.
- If X is discrete, $M_X(t) = \sum_{x \in \Omega'} e^{tx} p_X(x)$.
- If X is continuous, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$.
- For $Exp(\lambda)$ variable, $M_X(t) = \frac{\lambda}{\lambda - t}$ for $\lambda < t$.

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.
- If X is discrete, $M_X(t) = \sum_{x \in \Omega'} e^{tx} p_X(x)$.
- If X is continuous, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$.
- For $Exp(\lambda)$ variable, $M_X(t) = \frac{\lambda}{\lambda - t}$ for $\lambda < t$.
- Define $D_X := \{t : M_X(t) < \infty\}$. D_X is called the region of convergence (ROC).

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.
- If X is discrete, $M_X(t) = \sum_{x \in \Omega'} e^{tx} p_X(x)$.
- If X is continuous, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$.
- For $Exp(\lambda)$ variable, $M_X(t) = \frac{\lambda}{\lambda - t}$ for $\lambda < t$.
- Define $D_X := \{t : M_X(t) < \infty\}$. D_X is called the region of convergence (ROC).$t = 0$ is always part of ROC.

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.
- If X is discrete, $M_X(t) = \sum_{x \in \Omega'} e^{tx} p_X(x)$.
- If X is continuous, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$.
- For $Exp(\lambda)$ variable, $M_X(t) = \frac{\lambda}{\lambda - t}$ for $\lambda < t$.
- Define $D_X := \{t : M_X(t) < \infty\}$. D_X is called the region of convergence (ROC).$t = 0$ is always part of ROC.
- HW: Find the MGF for a random variable X that has the following distributions: Binomial(n,p), Normal $\mathcal{N}(0, 1)$, Poisson(λ)

MGF

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)
- Let ${M_X}^{(r)}(t) := \frac{d^r}{dt^r} M_X(t)$ (r^{th}-derivative of $M_X(t)$)

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)
- Let $M_X{}^{(r)}(t) := \frac{d^r}{dt^r} M_X(t)$ (r^{th}-derivative of $M_X(t)$)
- It can be shown that $M_X{}^{(r)}(t) = E[e^{tX} X^r]$ for all r and $|t| \leq \epsilon$.

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)
- Let $M_X{}^{(r)}(t) := \frac{d^r}{dt^r} M_X(t)$ (r^{th}-derivative of $M_X(t)$)
- It can be shown that $M_X{}^{(r)}(t) = E[e^{tX} X^r]$ for all r and $|t| \leq \epsilon$.
- $E[X^r] = M_X{}^{(r)}(0)$

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)
- Let ${M_X}^{(r)}(t) := \frac{d^r}{dt^r} M_X(t)$ (r^{th}-derivative of $M_X(t)$)
- It can be shown that ${M_X}^{(r)}(t) = E[e^{tX} X^r]$ for all r and $|t| \leq \epsilon$.
- $E[X^r] = {M_X}^{(r)}(0)$
- For $Exp(\lambda)$, ${M_X}^{(1)}(t) = \frac{\lambda}{(\lambda - t)^2}$ and ${M_X}^{(1)}(0) = E[X] = \frac{1}{\lambda}$

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)
- Let ${M_X}^{(r)}(t) := \frac{d^r}{dt^r} M_X(t)$ (r^{th}-derivative of $M_X(t)$)
- It can be shown that ${M_X}^{(r)}(t) = E[e^{tX} X^r]$ for all r and $|t| \leq \epsilon$.
- $E[X^r] = {M_X}^{(r)}(0)$
- For $Exp(\lambda)$, ${M_X}^{(1)}(t) = \frac{\lambda}{(\lambda - t)^2}$ and ${M_X}^{(1)}(0) = E[X] = \frac{1}{\lambda}$
- ${M_X}^{(r)}(0) = \frac{r!}{\lambda^r}$

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)
- Let ${M_X}^{(r)}(t) := \frac{d^r}{dt^r} M_X(t)$ (r^{th}-derivative of $M_X(t)$)
- It can be shown that ${M_X}^{(r)}(t) = E[e^{tX} X^r]$ for all r and $|t| \leq \epsilon$.
- $E[X^r] = {M_X}^{(r)}(0)$
- For $Exp(\lambda)$, ${M_X}^{(1)}(t) = \frac{\lambda}{(\lambda - t)^2}$ and ${M_X}^{(1)}(0) = E[X] = \frac{1}{\lambda}$
- ${M_X}^{(r)}(0) = \frac{r!}{\lambda^r}$
- HW: Find MGF of all random variables seen till now and use it to obtain moments.