Mathematical Statistics 1

Chapter 3. Some Special Distributions

3.4. The Normal Distribution—Proofs of Theorems

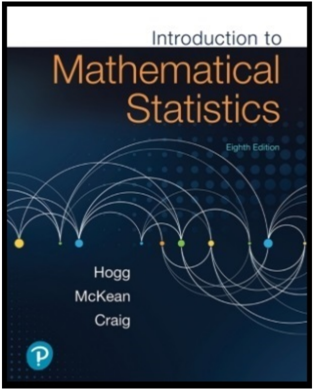

Table of contents

Theorem 3.4.1

Theorem 3.4.1. If the random variable X is $N(\mu, \sigma^2)$, where $\sigma^2 > 0$, then the random variable $V = (X - \mu)^2/\sigma^2$ is $\chi^2(1)$.

Proof. Let $W = (X - \mu)/\sigma$. Then W is $N(0,1)$ and $V = W^2$. The cumulative distribution function $G(v)$ of random variable V is, for $v \geq 0$,

$$G(v) = P(W^2 \leq v) = P(-\sqrt{v} \leq W \leq \sqrt{v}).$$

Since W is $N(0,1)$, then

$$\begin{aligned} G(v) &= \int_{-\sqrt{v}}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw \text{ for } v \geq 0 \\ &= 2\int_{0}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw \text{ for } v \geq 0 \end{aligned}$$

and $G(v) = 0$ for $v < 0$.

Theorem 3.4.1

Theorem 3.4.1. If the random variable X is $N(\mu, \sigma^2)$, where $\sigma^2 > 0$, then the random variable $V = (X - \mu)^2/\sigma^2$ is $\chi^2(1)$.

Proof. Let $W = (X - \mu)/\sigma$. Then W is $N(0, 1)$ and $V = W^2$. The cumulative distribution function $G(v)$ of random variable V is, for $v \geq 0$,

$$G(v) = P(W^2 \leq v) = P(-\sqrt{v} \leq W \leq \sqrt{v}).$$

Since W is $N(0, 1)$, then

$$\begin{aligned} G(v) &= \int_{-\sqrt{v}}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw \text{ for } v \geq 0 \\ &= 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw \text{ for } v \geq 0 \end{aligned}$$

and $G(v) = 0$ for $v < 0$. With the change of variables $w = \sqrt{y}$ then

$$G(v) = 2\int_0^{v} \frac{1}{\sqrt{2y}} e^{-y/2} \frac{1}{2} y^{-1/2}\, dy = \int_0^{v} \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-y/2}\, dy \text{ for } v \geq 0.$$

Theorem 3.4.1

Theorem 3.4.1. If the random variable X is $N(\mu, \sigma^2)$, where $\sigma^2 > 0$, then the random variable $V = (X - \mu)^2/\sigma^2$ is $\chi^2(1)$.

Proof. Let $W = (X - \mu)/\sigma$. Then W is $N(0, 1)$ and $V = W^2$. The cumulative distribution function $G(v)$ of random variable V is, for $v \geq 0$,

$$G(v) = P(W^2 \leq v) = P(-\sqrt{v} \leq W \leq \sqrt{v}).$$

Since W is $N(0, 1)$, then

$$\begin{aligned} G(v) &= \int_{-\sqrt{v}}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw \text{ for } v \geq 0 \\ &= 2\int_{0}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw \text{ for } v \geq 0 \end{aligned}$$

and $G(v) = 0$ for $v < 0$. With the change of variables $w = \sqrt{y}$ then

$$G(v) = 2\int_0^v \frac{1}{\sqrt{2y}} e^{-y/2} \frac{1}{2} y^{-1/2}\, dy = \int_0^v \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-y/2}\, dy \text{ for } v \geq 0.$$

Theorem 3.4.1 (continued)

Theorem 3.4.1. If the random variable X is $N(\mu, \sigma^2)$, where $\sigma^2 > 0$, then the random variable $V = (X-\mu)^2/\sigma^2$ is $\chi^2(1)$.

Proof(continued). ...

$$G(v) = \int_0^v \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-y/2}\, dy \text{ for } v \geq 0.$$

So the probability density function $g(v)$ is (rewriting in a suggestive way)

$$g(v) = \begin{cases} \frac{1}{\sqrt{\pi}2^{1/2}} v^{1/2-1} e^{-v/2} & \text{for } v > 0 \\ 0 & \text{for } v \leq 0. \end{cases}$$

Since $\Gamma(1/2) = \sqrt{\pi}$ then $g(v)$ is the chi-square distribution with $r = 1$, $\chi^2(1)$, as claimed. □

Theorem 3.4.1 (continued)

Theorem 3.4.1. If the random variable X is $N(\mu, \sigma^2)$, where $\sigma^2 > 0$, then the random variable $V = (X - \mu)^2/\sigma^2$ is $\chi^2(1)$.

Proof(continued). ...

$$G(v) = \int_0^v \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-y/2}\, dy \text{ for } v \geq 0.$$

So the probability density function $g(v)$ is (rewriting in a suggestive way)

$$g(v) = \begin{cases} \frac{1}{\sqrt{\pi}2^{1/2}} v^{1/2-1} e^{-v/2} & \text{for } v > 0 \\ 0 & \text{for } v \leq 0. \end{cases}$$

Since $\Gamma(1/2) = \sqrt{\pi}$ then $g(v)$ is the chi-square distribution with $r = 1$, $\chi^2(1)$, as claimed. □

Theorem 3.4.2

Theorem 3.4.2. Let $X_1, X_2, \ldots, X_n$ be independent random variables such that, for $i = 1, 2, \ldots n$, X_i has a $N(\mu_i, \sigma_i^2)$ distribution. Let $Y = \sum_{i=1}^n a_i X_i$ where $a_1, a_2, \ldots, a_n$ are constants. Then the distribution of Y is $N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_1^2\right)$.

Proof. The moment generating function of X_i is $M_{X_i}(t) = e^{\mu_i t + t^2 \sigma_i^2/2}$ for $t \in \mathbb{R}$ (as shown above), so by Theorem 2.6.1 the moment generating function of Y is

$$M_Y(t) = \prod_{i=1}^n M_i(t) = \prod_{i=1}^n e^{\mu_i t + t^2 \sigma_i^2/2} = \exp\left(\sum_{i=1}^n (\mu_i t + t^2 \sigma_i^2/2)\right)$$

$$= \exp\left(t \sum_{i=1}^n \mu_i + (t^2/2) \sum_{i=1}^n \sigma_i^2\right).$$

This is the moment generating function of a the normal distribution $N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma^2\right)$, as claimed. □

Theorem 3.4.2

Theorem 3.4.2. Let $X_1, X_2, \ldots, X_n$ be independent random variables such that, for $i = 1, 2, \ldots n$, X_i has a $N(\mu_i, \sigma_i^2)$ distribution. Let $Y = \sum_{i=1}^n a_i X_i$ where $a_1, a_2, \ldots, a_n$ are constants. Then the distribution of Y is $N\left(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_1^2\right)$.

Proof. The moment generating function of X_i is $M_{X_i}(t) = e^{\mu_i t + t^2\sigma_i^2/2}$ for $t \in \mathbb{R}$ (as shown above), so by Theorem 2.6.1 the moment generating function of Y is

$$M_Y(t) = \prod_{i=1}^n M_i(t) = \prod_{i=1}^n e^{\mu_i t + t^2\sigma_i^2/2} = \exp\left(\sum_{i=1}^n (\mu_i t + t^2\sigma_i^2/2)\right)$$

$$= \exp\left(t\sum_{i=1}^n \mu_i + (t^2/2)\sum_{i=1}^n \sigma_i^2\right).$$

This is the moment generating function of a the normal distribution $N\left(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma^2\right)$, as claimed. □